

R; Voorhorst, W.G.B.; Eggen, R.I.L.; Geerling, A.C.M.; Plattew, C.; Siezen, R.J.; Vos, J. Biol. Chem. 271, 20426-20431, 1996
 A; Title: Isolation and characterization of the hyperthermostable serine protease, pyroly
 A; Reference number: Z20481; MUID:96355370; PMID:8702780
 A; Accession: T28159
 A; Status: preliminary; translated from GB/EMBL/DBJ
 A; Molecule type: DNA
 A; Residues: 1-1398 <VOO>
 A; Cross-references: EMBL:U55835; NID:G1556462; PID:G1556463; PIDN:AAB09761.1
 A; Experimental source: DSM3638
 A; Genes: pls
 C; Keywords: hydrolase; serine proteinase

Query Match 13.5%; Score 305; DB 2; Length 1398;
 Best Local Similarity 28.4%; Pred. No. 9.5e-12;
 Matches 129; Conservative 37; Mismatches 153; Indels 136; Gaps 15;

QY 21 GQQQVAVADTGLDTCGRNDS-----SMHEAFRGKITALYALGRNNANDPN-- 66
 DB 301 GNGYDIAYVDTLDYDFTDEVPLGQNVYDYAVSYVYVGLN--YVLAELI-----DPNGE 354

QY 67 -----GHGTHVAGSVLGNALN-----XG 84

DB 355 YAVFGWDGEGHGHVAGTVAGYDSNNDWDWLSMYSGEWEVSRLYGDYTNVTDTVQG 414

QY 85 MAPQANLVFQSIIDSSGGLGGLPSNLNTLFSQAWNAGARIHTNSWG--APVNGAYTANSR 142

DB 415 VAPCAQIMAIRVLRSDG--RGSMDIIEGNTVAATHGADVINSLGGNAPYLDGTDPSV 472

QY 143 QVDEYVRNNDMTVLFAGNEGNSGTISAPGAKNAITVGATE----- 185

DB 473 AVDELTEKYGWPEVIAAGNEGGINIVGSPGVATRAITVGAAAAPIVNGVYVVSQALGYPD 532

QY 186 ----NVRPFGSLADNPNIHAFSSRGATDGRIPKDVTPAGCTFFILSARSSSLAPDSFWA 241

DB 533 YGCFYFPATNV-----RIAPFSSRGPRIDEIKRNVVAPGYGYSSLPWIGGADP-- 585

QY 242 NYNSKYAYVGGTSMATPIVAGNVAQUREHFINKRGITPKSLIKAAIAGAT-----DV 295

DB 586 -----NKGTSMATPHVSGVVALISG--PKPEGIYVNFIDIKKVLSEGATWLEGDPVT 636

QY 296 GLGYPSGQDQGWGVTLDKSLNVAYNNEATATGQATYSFOAQAGKPKLSLVWTDAPG 355

DB 637 GQKYTELDQGHGLVNVTKSWEI-----LKAINGTTLPIVDHWADKSY 678

QY 356 STTASVTLVNDLVLITAPNG-----QKTVGN 382

DB 679 SDFAELGV-DVIRGLYARNSIPDIVWHKIVGD 712

RESULT 7
 G83753 subtilisin-type proteinase (EC 3.4.21.-) vpr precursor [similarity] - Bacillus haloduran
 C; Species: Bacillus halodurans
 C; Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C; Accession: G83753
 R; Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A; Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A; Reference number: A83650; MUID:20512582; PMID:11058132
 A; Accession: G83753
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-799 <STO>
 A; Cross-references: GB:AP001510; GB:BA000004; NID:G10173440; PIDN:BA04550.1; GSPDB:GN00
 A; Experimental source: strain C-125
 C; Genes: C
 C; Keywords: microbial serine proteinase vpr; subtilisin homology
 F; 1-29/Domain: signal sequence #status predicted <SIG>

Query Match 13.4%; Score 302.5; DB 2; Length 799;
 Best Local Similarity 24.0%; Pred. No. 6.5e-12;
 Matches 150; Conservative 57; Mismatches 152; Indels 265; Gaps 28;

QY 18 GLYGQGVAVADTGLDTCGRNDSMHAEFRKITALYALGR-----TNN-----AN 63

DB 171 GYTGEITVAIILDTGVDTYTHPD-----LVHAFGDYKGMDFIDNDDPQPTPG 218

QY 64 DPNQ-----HGTHVAGSVLGNALNKGMAPOANLVFQISGSSGGLGGLPSNLNTL--FSQA 117

DB 219 DPRGIETHTGTHVAGTVAANGLIKGVAPDANLLAYRL-----GPGRGSTAGVIAGIERA 274

QY 118 WNAGARIHTNSGAPVNGAYTANSROVDYVRNNDMTVLFAGNEGPNSTGTSAPGTAKN 177

DB 275 VQGDADIMNLGLNTLNDPDPFATSIADL-WMAAGVAVVTSNGSGPNWTVGSPGTSRD 333

QY 178 AITVGAT-----ENYRPSFGSL----- 194

DB 334 AISVGATRLPYNKYKASVFTSDGIDYPSADIWGFPSPDELLDGETYEFAGLKGPGD 393

QY 195 -----ADNP----- 198

DB 394 PEGVDVEGKALIVRGEIPFVEKAENAKAAGAVGAILIYNNVAGVQTPVGLAIPTIMLSN 453

QY 199 -----NHIAQFSSRG-ATRDGRIRKPDVTPAGTFFILSA 229

DB 454 EDGLKWRNELENGQNTVTFSEIPDKLVGETVADFSRGPVMTWMIKPDVSAFGVAIVST 513

QY 230 RSSLADPSSFWANYSKYVNGGTSMATPIVAGNVAQLREHFINKRGITPKSLIKAAI 289

DB 514 IPHQPDPPY--GYGRSQ-----GTSWASPHVAGAAALLLEAH-PWNGV-----DHVKAALM 562

QY 290 AGATDV-----GLGYPSGQDQGWGVTLDKSLNVAYNNEATALAT----- 328

DB 563 NTAENLVDENGNRYPHNTQAG-----SIRIVDA-ESETLVTPGSHSGFTTKERKQV 616

QY 329 -----GQATYSFOAQ-AGKP-----LKISLVTDAPGSTTASYTLVNDLIVIT 372

DB 617 EROHFTIHNSLRKTYQFQVQAGNPDGIKVKTSLNRVQPKT----- 661

QY 373 APNGQKQVGN-----DFSYPYDNN-----WDGRNVE-----NVFINA-----PQSGTYTIEV 415

DB 662 ----QKINENVQVDARKLDPGY-YEGTIIIVSDGSQTVVEPTILFVSEPDYPRVTTFDLDI 716

QY 416 -----QAYNVPSGQRFSLAI 431

DB 717 DENGVLFGSAY-LFNGAEFEGLMI 739

RESULT 8

S11890 serine proteinase (EC 3.4.21.-) precursor, extracellular - Xanthomonas campestris pv. c
 N; Alternate names: subtilisin-related proteinase
 C; Species: Xanthomonas campestris pv. campestris
 C; Date: 21-Nov-1993 #sequence_revision 07-Feb-1997 #text_change 03-Dec-1999
 C; Accession: S11890
 R; Liu, Y.N.; Tang, J.L.; Clarke, B.R.; Dow, J.M.; Daniels, M.J.

Mol. Gen. Genet. 220, 433-440, 1990
 A; Title: A multipurpose broad host range cloning vector and its use to characterise an
 A; Reference number: S11890; MUID:90251253; PMID:2187155
 A; Accession: S11890
 A; Molecule type: DNA
 A; Residues: 1-580 <LIU>
 A; Cross-references: EMBL:X51635; NID:948533; PIDN:CAA35962.1; PID:948534

A; Experimental source: Xanthomonas campestris pv. campestris
 A; Note: the sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 205-AL
 C; Superfamily: subtilisin, subtilisin homology
 C; Keywords: extracellular protein; hydrolase; serine proteinase
 F; 1-32/Domain: signal sequence #status predicted <SIG>
 F; 168-423/Domain: subtilisin homology <SBT>

Query Match 13.2%; Score 298.5; DB 2; Length 580;
 Best Local Similarity 29.0%; Pred. No. 7.6e-12;

Matches 137; Conservative 51; Mismatches 149; Indels 135; Gaps 26;

QY 21 GCGVAVVADTGL-----DTGRNDSMHEAFRGKITALVALGRTNAND----- 64

Db 158 GSGTVAVDITGITHSHADLNANILAGYDFISDAITARDGCRDSDNADEGWYAANFEGA 227

QY 65 -----PNHGTHVAGS-----VLGNLNKGMAPQANL-----VFQSIMDS 99

Db 228 GIPAAASSWHGTHVAGTVAATNTTGVAGTAYGAKVVPVRVLGKCGGSLSDIADAIWA 287

QY 100 SGG-LGGIPLSNLN--TLFSQAWNACARHTNSWAGVUNGAYTANSRQVDEYVRNNDMTVL 156

Db 288 SGGTVSGIPANANPAEVNIMSLGGGSCSTTQNA--INGAVSRGT-----TVV 334

QY 157 FAAGNEGN-SGTISAPGTAKNAITVGATENYRPSFGSLADNPXNHAOFSSRGATRDGR 215

Db 335 VAAAGNDASNVSG--SLPANCAVIAVAAT--TSAGAK-----ASYSNFGT----- 375

QY 216 KPDVTAPGTFTLSARSS--LAPDSSFMANYSKYAMGTSMAPIVAGNVAQLREHF 273

Db 376 GIDVASPGSSILSTUNGSTTTPGSAASYN-----GTSMASPHVAGVVALVQS--VA 426

QY 274 NRGITPK--PSLIK--AALIAGATDVGLGY-----PSGDQGWGRVTLDKS 314

Db 427 PTALTAAVETLTKNTARALPGACSGCGAGVWADAAVTAALNGSGGGGGGNTLNG 486

QY 315 LNVAVNEATATATQKATYSFOAGKPLKLSLWTDAPGSTTASYTLVNDLVL----- 371

Db 487 TPVT-----GLGAATGAELNYITVPAGSG--TLTVTTSGGS-----GDADLYVRAG 531

QY 372 TAPNQKQVGNDFSPYDNNDGRNVENFINAPQSGTYTIEVQAYNVPSG 423

Db 532 SAPTDSAYT---CRPYRS-----GNAETCITAP-SGTYYVRLKAYSTPFG 573

RESULT 9

T29090

Surface layer-associated STABLE proteinase - Staphylothermus marinus

N/Alternate names: hyperthermostable proteinase

C/Species: Staphylothermus marinus

C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000

C/Accession: T29090

R/Mayr, J.; Lupas, A.; Kellermann, J.; Eckerskorn, C.; Baumeister, W.; Peters, J. Curr. Biol. 6, 739-749, 1996

A/Title: A hyperthermostable protease of the subtilisin family bound to the surface layer

A/Reference number: Z20559; MUID:96385442; PMID:8793300

A/Accession: T29090

A/Status: Preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1345 <MAY>

A/Cross-references: EMBL:U57968; NID:gi374755; PID:gi374755; PIDN:AA02323.1

A/Experimental source: strain F1

C/Function:

A/Description: probably serves an exodigestive function related to the organism's energy

A/Note: stoichiometric S-layer component

Query Match 12.6%; Score 285; DB 2; Length 1345;

Best Local Similarity 26.1%; Pred. No. 1.8e-10;

Matches 123; Conservative 60; Mismatches 156; Indels 132; Gaps 19;

QY 46 FRGKITALVALGRTNANDPNCHGTHVA-----GSLGNALN-----KGMAPQANLV 92

Db 445 YQRYLAL-----VSPFHGHGTSVAVIASRGRVLDYLDGDKLYRMGVAPGAKI- 495

QY 93 FQSIMDSSGGLPLSNLTLFSQAWNAG-----ARHTNSW 129

Db 496 -----AGGDALLGNI--LVLEAWLAGFNIVTEEDGYVYLSLDPFGPHRADIIISNW 546

QY 130 GA-----PVCAYTANSRQVDEYVRNNDMTVLFAAGNEGNSGTISAPGTAK 176

Db 547 GSIYINFWLQOPFGIDYRSFMDLAIARNYLIQDHVTIVFAAGNEGPGYSNGAPGTGL 606

QY 177 NAITVGATE--NYRPSFGSLADNPXNHAOFSSRGATRDGRIPDVTAPGTFTLSARSSIA 234

Db 607 LVITAGASTLWDYTRIYGYPEGVADEVIPFSSRGPTGQGYPKPDIIVNIGAF----- 657

QY 235 PDSSFWAN-----YNSKYAVMGTSMAPIVAGNVAQLREHFKNRGITPKPSLIKA 286

Db 658 -----EWASRTTIDRGYGAQPDVFGTSEATPYTSGTLALVFOAYKEVNTTDPPTAKI 713

QY 287 ALIAGATDVGLGYPSGDQGWGRVTLDKSLN-----VAYVNEATATATGQKATYSFOAQ 339

Db 714 ILKSSAKDI--WYPAPSQSGRVDALKAADTVFISEWLAYVSEGIQAEFLENVYTDGPGY 771

QY 340 AKPLKLSLWTD-----APGSTTASYTL--VNDLDELVTAPNG---QKYVGNDSFSPYD 389

Db 772 IGVLVLPY-LADTDYGVVVRPGS--SKNFTLNIYVNGAVLSAWNTVLYKEITYDVGVDYS 829

QY 390 N-----NWDGRNVE-----NVFINAPQSGTYTIEVQAYN 419

Db 830 GLLFLKVPKYAGSADYVEVWQLENMTPPGVFLKTPIDPLHAILISAYD 880

RESULT 10

JC4908

alkaline serine proteinase (EC 3.4.-.-) I precursor - Alteromonas sp.

N/Alternate names: sutillase

C/Species: Alteromonas sp.

C/Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 08-Oct-1999

C/Accession: JC4908

R/Tsuji, H.; Miyamoto, K.; Tanaka, K.; Kaideu, Y.; Imada, C.; Okami, Y.; Inamori, Y. Biosci. Biotechnol. 60, 1284-1288, 1996

A/Title: Cloning and sequence analysis of a protease-encoding gene from the marine bacterium *Alteromonas*

A/Reference number: JC4908; MUID:97141200; PMID:8987544

A/Accession: JC4908

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-715 <TSU>

A/Cross-references: DDBJ:D38600; NID:gi536787; PIDN:BAAL8912.1; PID:dl019647; PID:92160

A/Experimental source: strain O-7

C/Comment: This enzyme belongs to class I subtilisin-like family. It is a chelator-sens

C/Genetics:

A/Gene: aprI

A/Superfamily: subtilisin homology

C/Keywords: hydrolase

F/1-40/Domain: signal sequence #status predicted <SIG>

F/41-150/Domain: amino-terminal propeptide #status predicted <ATP>

F/151-496/Product: alkaline serine protease I #status predicted <MAT>

F/182-452/Domain: subtilisin homology <SBT>

F/457-715/Domain: carboxyl-terminal propeptide #status predicted <CTP>

F/239-294,335-372,478-481/Disulfide bonds: #status predicted

Query Match 12.1%; Score 272.5; DB 2; Length 715;

Best Local Similarity 25.8%; Pred. No. 4.9e-10;

Matches 126; Conservative 43; Mismatches 164; Indels 155; Gaps 20;

QY 21 GCGVAVVADTGL-----LDTGRNDSMHEAFRGKITALVALGRTNANDP----- 65

Db 182 GCGVAVVADTGYRPHLLDANILPDYDMISNTFVNDGGARDNDARDGDVATRGCGT 241

QY 66 -----NGHGTHTVAGSVLGNALN-----KGMAPQANLVFQSIMDSSGGLGGLP 107

Db 242 DSSGQPVPRADQDSSWHGTHVAGTVAATNNGEGVAGVADAKVVPVRVL---GKCGGLT 298

QY 108 SNLTLFQAWNAGARIHTNSWGPV-----NGAYTANSRQVDEYVRNNDMTVLFAAG 160

Db 299 SDIADGIIWAGSGSDRPANPANPAVNNMISLGGGACSAATTQAINQARNNGTVIVIAAG 358

QY 161 NEGPSQGTISAPGTAKNAITVGATENYRPSFGSLADNPXNHAOFSSRGATRDGR----- 214

Db 359 NDNDNSA-----VYNPG-----NCNGVNVASVG--RDGSRAYSN 392

QY 215 --TKPDVTAPGTFTLSARSSLPD-----SSFWANYSKYAMGTSMAPIVAGN 263

Db 393 YGANIDVAAFG-----GAQFADDPGEGILSTHNSGSGAPSDNSYHSQGTSMAPHVAGV 447

| | | | |
|----|-----|--|------|
| QY | 264 | VAQRREHFIKNRGITPKPSLIKAALAGATDVGLGVPSCDQGGRVTLDKSLNVAYNEA | 323 |
| | : | ::: | :::: |
| Db | 448 | AALIQQ-----AKPSTAPDEVEFILKNTTRSFAGSCNSCGTVGDAA---AAVNEA | 495 |
| QY | 324 | -----TAL--ATCQKATYSFQAQAGKPLKISLWTDAPGSTTASYT | 362 |
| Db | 496 | LGDVVPTPTCNTLEDGVAKTGLSGAAGSNQFFTFDPVPAK-----TNVTFT | 541 |
| QY | 363 | L----VNVDLVIIAPNGOKVGNDFSPYPNNWDGR-----NVENVFINAPQSQTIVIEV | 415 |
| Db | 542 | MSGGTGDADLYVK-----LG-----SQTSSSYDCRPEYEGNAEVCSDFAQAGTYHVI | 591 |
| QY | 416 | QAYNVPSG | 423 |
| Db | 592 | NGYKAYSG | 599 |

RESULT 11

A49778
high-alkaline serine proteinase (BC 3.4.21.-) precursor - Bacillus alcalophilus (strain N; Alternate names: subtilisin homolog, high-alkaline C; Species: Bacillus alcalophilus C; Date: 12-Mar-1994 #sequence_revision 24-Feb-1995 #text_change 20-Jun-2000 C; Accession: A49778; JCI244
R; van der Laan, J.C.; Gerritse, G.; Mulleners, L.J.S.M.; van der Hoek, R.A.C.; Quax, W.J.
Appl. Environ. Microbiol. 57, 901-909, 1991
A; Title: Cloning, characterization, and multiple chromosomal integration of a Bacillus A; Reference number: A49778; MUID: 91282483; PMID: 2059048
A; Accession: A49778
A; Molecule type: DNA
A; Residues: 1-380 <VAN>
A; Cross-references: GB:M5086; NID:G142456; PIDN:AAA22212.1; PID:G142457
A; Experimental source: strain PB92, ATCC 31408
A; Note: amino end of mature protein confirmed by peptide sequencing
R; Takami, H.; Kobayashi, T.; Kobayashi, M.; Yamamoto, M.; Nakamura, S.; Isono, R.; Horike
Biochim. Biotechnol. Biochem. 56, 1453-1460, 1992
A; Title: Molecular cloning, nucleotide sequence, and expression of the structural gene B; Reference number: JCI244; MUID: 93043753; PMID: 1368952
A; Accession: JCI244
A; Molecule type: DNA
A; Residues: 1-195, 'S', 197-380 <TAK>
A; Cross-references: GB:P13157; NID:G216231; PIDN:BAAC2442.1; PID:G216232
A; Experimental source: Bacillus alcalophilus Vedder, ATCC 21522 (Bacillus sp. 221)
C; Superfamily: subtilisin, subtilisin homology
C; Keywords: hydrolase; serine proteinase; zymogen
F; 1-27/Domain: signal sequence #status predicted <SIG>
F; 28-111/Domain: activation peptide #status predicted <PRO>
F; 112-380/Product: alkaline serine proteinase #status predicted <MAT>
F; 134-340/Domain: subtilisin homology <SBT>
F; 143, 173, 326/Active site: Asp, His, Ser #status predicted

Query Match 11.9%; Score 269; DB 2; Length 380;
Best Local Similarity 31.8%; Pred. No. 3.5e-10;
Matches 97; Conservative 35; Mismatches 99; Indels 74; Gaps 14;

| | | | |
|----|-----|---|------|
| QY | 13 | AQNNGILYGQGVAVADTGDTGRNDSSMHEAFRGKITALVALGRNTNANDPNGHGTHV | 72 |
| Db | 126 | AAHNRLGTGSQKVAVLDTGIST-----HPDLNIRGGASFVPGEP-STQDNGHGHTV | 177 |
| QY | 73 | AGSV--LCNALN-KGMAPQANILFQSDMSGGGLGPSNLNTLFSAQWAGARIHTNSW | 129 |
| Db | 178 | AGTIAALNNIGVLGVAPNLAELYAVKVLGASG--SGSVSSIAQGLEWAGNKGHVANLSL | 235 |
| QY | 130 | GAPVNGAYTANSROVDYVRNMDTVLFAAGNEGPNSTGISAPCTAKNAVITVGATENYRP | 189 |
| Db | 236 | GSP---SFSATLEQAVNSATSRSGLVVAASGNSG--AGSISYPARYANAMAVGAT----- | 285 |
| QY | 190 | SFGSLADNPNHIAQFSRSGATRDRGRIIPDYTPACTFTILSRSSLAPDSSFWANNYSKYAY | 249 |
| Db | 286 | -----DONNRRASFSQVAGL-----DIVAPGVNVOSTYP-----GSTYAS | 321 |
| QY | 250 | MGGTSMATPIVAGNVA-----OLREHFINKRGITPKPSLIKAALAGATDVGL | 297 |
| | : | ::: | :::: |

```

Db      322  LNCSTATHVAGAAALVKKQKPSWNVQIRNH-LKN-----TATSLGSTNL-- 367

Qy      298  GYPESG 302
      368  -YGSIG 371

Db

RESULT 12
JS0173
A:alkaline proteinase (EC 3.4.21.-) A precursor - Vibrio alginolyticus
N:Alternate names: alkaline serine exoproteinase A
C:Species: Vibrio alginolyticus
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C/Accession: JS0173
R/Deane, S.M.; Robb, F.T.; Robb, S.M.; Woods, D.R.
Gene 76, 281-288, 1989
A/Title: Nucleotide sequence of the Vibrio alginolyticus calcium-dependent, detergent-r-
A/Reference number: JS0173; MUID:89326126; PMID:2546851
A/Accession: JS0173
A:Molecule type: DNA
A:Residues: 1-534 <DEA>
A/Cross-references: GB:M25499; NID:G155250; PID:AAA27550.1; PID:G155251
C/Comment: This protein is a calcium-dependent and sodium dodecyl sulfate-resistant pro
C/Superfamily: subtilisin; subtilisin homology
C/Keywords: calcium; hydrolase; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-534/Product: alkaline serine proteinase A #status predicted <MAT>
F:171-378/Domain: subtilisin homology <SBT>
F:271-273,297-299/Region: SI specificity crevice #status predicted
F:180,213,363/Active site: Asp, His, Ser #status predicted

Query Match      11.7%; Score 265; DB 1; Length 534;
Best Local Similarity 26.3%; Pred. No. 1e-09;
Matches 123; Conservative

Qy      7  IVKADVAQ-----NNY--GLYGGQVVAVADTGLDTRGNDSSKHEAF 46
Db     137  IVSADANQTNAIWGLDRIDORNFLDNINSGANFDGTGVTAYVIDTG-----NNAHVEF 190
Qy     47  RGKITALY-ALGRNTNANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIM--DSSGGL 103
Db     191  GGRSVSGYDFVDNDADNSCNGHGHVAGTIGSGLY--GVAKVNLVGVTVLSCSGSGST 248
Qy     104  GGLPSNLTLFSAQWAGARIHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEG 163
Db     249  SGVIAGVDWAANA--SGPSVANWSLG---GGQSVALDSAVQSAVQSG-VSFMLAAGNSN 302
Qy     164  PMSGTISAPGTAKNAITVGATENYRSPFGLADPNPHIAQFSSRGATRDGRIKPDVTAPG 223
Db     303  ADACNTS-PAKVATGTVTGST-----TSDARSSFSNWGSC-----VDVFAPG 344
Qy     224  TFLSARSSLAPDSSFWANNYSKYAYMGTSMATPIVAGNVAQLREHFIKNRGITTPFSL 283
Db     345  SQIXSA-----W--YDGGYKITSGMTATPHVAG-VAAL--VLQENSSVS--PSQ 387
Qy     284  IKAALTAGA-----TVGLGLYPSG-----DQWGRVTLDKSLNVAY 319
Db     388  VEALIVSRASGTGKVTDRGSVKNLLYSLTDADCGQDCGPDPTDPDPGKLTSGVPVS--- 444
Qy     320  VNEATALATGKATYSFQAQAGPLKISLVWTDAPGSTTASYTLVNDLDELVITAPNGOKY 379
Db     445  ---GLSGSSGQVAYYYVDVBAQRLTVQM-----YGGSGDADLYLRF--GAK- 486
Qy     380  VGNDFSFYFDNNWDGR-----NNVENFINAPQSGTYTIEVOAYNVPESG 423
Db     487  -----PTINAMDRCPFYKGNNETCTVSTQSGRYHVMICGYSNYSG 527

```

RESULT 13
S23407
subtilisin (EC 3.4.21.62) 1 precursor - *Bacillus* sp. (strain TA39)
C:Species: *Bacillus* sp.
C:Date: 04-Dec-1992 #sequence revision 04-Dec-1992 #text change 18

C;Species: Bacillus sp.
C;Date: 04-Dec-1992 #sequence revision 04-Dec-1992 #text change 18-Jun-1999

C;Accession: S23407
R;Narinx, E.; Davail, S.; Feller, G.; Gerday, C.
Biochim. Biophys. Acta 1131, 111-113, 1992
A;Title: Nucleotide and derived amino acid sequence of the subtilisin from the antarctic
A;Reference number: S23407; MUID:92256481; PMID:1581352
A;Accession: S23407
A;Molecule type: DNA
A;Residues: 1-420 <NAR>
A;Cross-references: EMBL:X62369; NID:g40200; PIDN:CAA44227.1; PID:g40201
C;Genetics:
A;Gene: sub1
C;Superfamily: subtilisin; subtilisin homology
C;Keywords: extracellular protein; hydrolase; serine proteinase
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-111/Domain: propeptide #status predicted <PRO>
F;112-420/Product: microbial serine proteinase #status predicted <MAT>
F;136-374/Domain: subtilisin homology <SBT>
F;145,185,360/Active site: Asp, His, Ser #status predicted

Query Match 11.7%; Score 263.5; DB 1; Length 420;
Best Local Similarity 29.0%; Pred. No. 9.2e-10;
Matches 92; Conservative 46; Mismatches 120; Indels 59; Gaps 15;

Qy 21 GGGQVAVADTGLTGRNDSMH-EAPRGKITAYALGRT---YALGRT---NNANDPNHGTHVAGS 75
Db 136 GGGINIAVLDTGVNTN-----HPDLNNVEQCKDFVTGTYTNNSCDRQGHGTHVAGS 189

Qy 76 VL-----GNALNGMAPOANLVFQSIM--DSSGGLGGLPSNLNTLFSQAWNAGAR.IHTN- 127
Db 190 ALADGGTNGGV-YGVAPDADLWAYKVLGGDGGYADDIAAAIRHAGDQATALTNTKVVINM 248

Qy 128 SWGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNENSGTISAPGTAKNAITVGATENY 187
Db 249 SLGSSGESSLITNA--VNYSYNGKVLIIAAAGNSGPGYQSGISYPGALVNAVAAALEN- 304

Qy 188 RPSFGLADPNHIAQFSSEG-----ATDRGRKPDVTPAPCTFILSARSLAPDSF 239
Db 305 -----KVENTYVADFSSISGYSWTGDAIAIQGDV--EISAPGAAYST----- 347

Qy 240 WANYNSKYAYMGTSVATPIVAGNVAQLREHFKNKGITPKPSLIKAALIAGATDVGLCY 299
Db 348 W--FDGGYATISGTSVAPHAAGLAAKINAQYPSASNVDRGELQYRAY---ENDILSGY 402

Qy 300 PSG-----DQGWREVLT 311
Db 403 YAGYGDGDFASGFGFATV 419

RESULT 14
S25835
subtilisin (EC 3.4.21.62) precursor - Bacillus sp. (strain TA41)
C;Species: Bacillus sp.
C;Date: 22-Nov-1993 #sequence_revision 20-Feb-1995 #text_change 18-Jun-1999
C;Accession: S25835
R;Davail, S.; Feller, G.; Narinx, E.; Gerday, C.
Gene 119, 143-144, 1992
A;Title: Sequence of the subtilisin-encoding gene from an antarctic psychrotroph Bacillus
A;Reference number: S25835; MUID:93012966; PMID:1398082
A;Accession: S25835
A;Molecule type: DNA
A;Residues: 1-419 <DAV>
A;Cross-references: EMBL:X63533; NID:g40198; PIDN:CAA45096.1; PID:g40199
C;Superfamily: subtilisin; subtilisin homology
C;Keywords: extracellular protein; hydrolase; serine proteinase
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-110/Domain: propeptide #status predicted <PRO>
F;111-419/Product: microbial serine proteinase #status predicted <MAT>
F;135-373/Domain: subtilisin homology <SBT>
F;144,184,359/Active site: Asp, His, Ser #status predicted

Query Match 11.3%; Score 256.5; DB 1; Length 419;
Best Local Similarity 30.7%; Pred. No. 2.6e-09;
Matches 81; Conservative 41; Mismatches 99; Indels 43; Gaps 12;

Qy 21 GGGQVAVADTGLTGRNDSMH-EAPRGKITAYALGRT---NNANDPNHGTHVAGS 76
Db 135 GAGINIAVLDTGVNTNHPDLNNVEQCKD---FTVGTNFTNSCTDRQGHGTHVAGSA 189

Qy 77 LGNALN-----KGMAPOANLVFQSIM--DSSGGLGGLPSNLNTLFSQAWNAGAR.IHTN-SW 129
Db 190 LANGGTGSVGVGAPEADLWAYKVLGGDGGYADDIAAIRHAGDQATALTNTKVVINSL 249

Qy 130 GAPVNGAYTANSRQVDEYVRNNDMTVLFAGNENSGTISAPGTAKNAITVGATENYRP 189
Db 250 GSSGESSLITNA--VD-YAYDKGVLIIAAAGNSGPKFGSIGYPGALVNAVAAALEN--- 303

Qy 190 SFGSLADPNHIAQFSRGRTRDG-----RIKPDVTPAGTFTLSARSLAPDSFPWANY 243
Db 304 ---TIQGTVRVADFSSRGRHKRTAGDVYIQKGDVEISAPGAAYST-----W--F 348

Qy 244 NSKYAYMGTSVATPIVAGNVAQL 267
Db 349 DGGYATISGTSVAPHAAGLAAKI 372

RESULT 15
A72647
probable surface layer-associated STABLE proteinase APE0607 - Aeropyrum pernix (strain
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: A72647
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Taka
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: A72647
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1331 <KAW>
A;Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA79577.1; PID:d1043363; PID:g51
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0607

Query Match 11.3%; Score 256.5; DB 2; Length 1331;
Best Local Similarity 23.9%; Pred. No. 1.2e-08;
Matches 112; Conservative 72; Mismatches 153; Indels 131; Gaps 20;

Qy 27 AVADTGLTGRNDSMH-EAPRGKITAYLA-----DPNG-----HGTHVAGSVLGNALN----- 82
Db 360 ALSDTGMISGEPPDPSLLDLSPADFTPASVGEVLARDFTGCVNDFSGALAGWTYDWVG 419

Qy 56 --LQRTNNAN-----DPNG-----HGTHVAGSVLGNALN----- 82
Db 420 LLTGESVNLGWRGLGFDYAGLVLPGLDPQGRWVSILYDTLAHGTSA-TVTSRGNVEFNL 478

Qy 83 -----KGMAPOANLVFQSIMDSSGG-----LGGL---PSNLNTLFSQAWNA 120
Db 479 GYIETSIRGVAPGAKI-----AAGSFLINVFVAQLFLSGPEQDPSLNNWYTGHEOV 531

Qy 121 GAR.IHTNSWG---APVNGAYTA--NSRQVDEYVRNNDMTVL-FAAGNENSGTISAPGT 174
Db 532 D--VINNSWNGSYIALRGFLTGADYATIEDYIVSASGTIVHAMGNGGPGYGTATTFGA 589

Qy 175 AKNAITVGATE--NVRPSFGSLADPNHIAQFSRGRTRDGR.IKFDVTPAGTFTLSARSS 232
Db 590 GSLLISVQASTLFDYRPFYGVLPSPGGDVISWDRGPSQIGVAKPDVNVIGSPFAWAG--- 646

Qy 233 LAPDSSFWANYNSKYAYMGTSVATPIVAGNVAQLREHFKNKGITPKPSLIKAALIAGA 292
Db 647 -VPVLTGLNGSLAPDIFFGTSEATPMTSGSVALVSAVYQQAQAFGAKSPGLVKALKSTA 705

Qy 293 TDVGLGVPSPGQDQGRVTLDK-----SLNAVYVNEATALTATGQATYSPQAQA 340
Db 706 RDTGA--DAFTQSGSQGVDDVYRAVKAVLEGGVPITALSTSVENVYSLSSG----YSYFPLA 759

Search completed: March 31, 2004, 16:10:20
Job time : 12.3246 secs

283 LeuIleLysAlaAlaLeu-----IleAlaGlyAlaThrAspValGlyLeuGlyTyr 299
Db TGGACCTCGCGCGAACTGAAGCGCGCTCACCGGCTCCACCAAGGGCGGC---AAGTAC 1777
300 ProAsnGlyAspGlnGlyTrpGlyArgValThrLeuAsnLysSerLeuAsnValAlaTyr 319
Db ACCCGTTCGAGCAGGGTTCGGCGCGATCCAGGCGGACAAAGCGCTCCAGCAGACCGTG 1837
320 ValAsnGluAlaThrAlaLeuAlaThrGly----- 329
Db ATCGCGACCGCGTCTCGGTGAGTTCGGCGTCCAGCAGTGGCGGCACACCGACGACGAG 1897
330 -----GlnLysAlaThrTyrSerPheGlnAlaGlnAlaGlyLysProLeuLysIle 346
Db CCGGTCAACAGCAGTGCACCTACCCCACTCCGACCCAGCAGCAGCGTCAAGTGAAGTG 1957
347 SerLeuValTrpThrAspAlaProGlySerThrThrAlaSer-----TyrThrLeuVal 364
Db ACGTCGACCGCACCCAGCACCCCAAGGGCAAGGGCGGCCCGCGGCTTCTTCACGCTGGGC 2017
365 AsnAspLeuAspLeuValIleThrAlaProAsnGlyGlnLysTyrValGlyAsnAspPhe 384
Db -----GCCACCAACGAGTACCGTCCCGCGGGCGGC----- 2047
385 SerTyrProTyrAspAsnAsnTrpAspGlyArg-----AsnAsnValGlu 399
Db AGCGCTCCGTGCACATGACCGCGCGACACCGCGTCCGCGGACGCGTGGACGCGGTAC 2107
400 AsnValPheIleAsnAlaProGlnSerGlyThrTyrIle----- 412
Db TCGCGGTACGTGTGTCGCACCGCGCGGCGGCGACAGCGTCCGCACGCGCGCGGTGCAG 2167
413 IleGluValGlnAlaTyrAsnVal-ProSerGly 423
Db CGCGAGTGCAGTGCATGACGTCGACGTCGCGTCCGCG 2201

RESULT 14
US-09-000-016-1
; Sequence 1, Application US/09000016
; Patent No. 6143541
; GENERAL INFORMATION:
; APPLICANT: AKIRA ARISAWA et al.
; TITLE OF INVENTION: GENE ENCODING A PROTEIN HAVING ASYMMETRIC
; TITLE OF INVENTION: HYDROLASE ACTIVITY FOR 4-SUBSTITUTED 1,4-DIHYDROPYRIDINE DERIV
; TITLE OF INVENTION: ITS EXPRESSION PRODUCT
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., #800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/000,016
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; TELEX:

QY 358 ThrAlaSerTyThrLeuValAsnAspLeuAspLeuValIleThrAlaProAsnGlyGln 377
Db 1423 -----ACGGGCTCGAGCGACATCGACCTCTACTCTACGACCCCAAC----- 1464
QY 378 LysTyrValGlyAsnAspPheSerTyProTyrAspAsnAsnTyrAspGlyArgAsnAsn 397
Db 1465 -----GGGACGAGGTTGACTACTCTCTACCGCTACTAC-----CGC 1503
QY 398 ValGluAsnValPheIleAsnAlaProGlnSerGlyThrTyriIleGluValGlnAla 417
Db 1504 TTTCAGAGGTCGGCTACTACACCGCGGGAACCTGACGCTCAAGGTCGTCTCAGC 1563
QY 418 TyrAsnValProSerGlyProGlnArgPheSerLeuAlaIleVal 432
Db 1564 TAC-----AAGGGCGCGGGAACCTACAGGTCGAGCTGCTC 1599

RESULT 12
US-09-000-016-3
; Sequence 3, Application US/09000016
; Patent No. 6143541
; GENERAL INFORMATION:
; APPLICANT: Akira ARISAWA et al.
; TITLE OF INVENTION: GENE ENCODING A PROTEIN HAVING ASYMMETRIC
; TITLE OF INVENTION: HYDROLASE ACTIVITY FOR 4-SUBSTITUTED 1,4-DIHYDROPYRIDINE DERIV
; TITLE OF INVENTION: ITS EXPRESSION PRODUCT
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, L.L.P.
; STREET: 2033 K Street, N.W., #800
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/000,016
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-721-8200
; TELEFAX: 202-721-8250
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Streptomyces viridosporus
; STRAIN: A-914
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 338...2539
; IDENTIFICATION METHOD: E
US-09-000-016-3
Alignment Scores: 9,19e-25 Length: 2539
Pred. No.: 345.50 Matches: 132
Score:

Percent Similarity: 41.15% Conservative: 54
Best Local Similarity: 29.20% Mismatches: 170
Query Match: 15.29% Indels: 96
DB: 3 Gaps: 16
US-09-985-689A-4 (1-433) x US-09-000-016-3 (1-2539)
QY 18 GlyLeuTyrGlyGlnGlnValValAlaAlaAspThrGlyLeuAspThrGlyArg 37
Db 998 GGTACGACGGCAAGGGGTGAGATCGCGCTCTCGACACCGGTGTGCACACGAGC--- 1054
QY 38 AsnAspSerSerMetHisGluAlaPheArgGlyLysIleThrAlaLeuTyrAlaLeuGly 57
Db 1055 -----CATCGGACCTCAAGGGCGGGTCCCGCTCCCAAGAACTTCACC 1099
QY 58 ArgThrAsnAsnAlaSerAspProAsnGlyHisGlyThrHisValAlaGlySerValLeu 77
Db 1100 GCGCGCGCGCGCGCGGACAGGTGGCCAGCGACCCACCGTCCCTCGATCGCGGG 1159
QY 78 GlyAsnAlaLeuAsn-----LysGlyMetAlaProGlnAlaLeuVal 92
Db 1160 GGCACGGCGCGCCAGTCCAAAGGCAAGTACAGGGCGGTGCGACCCCGCGCGATCCTC 1219
QY 93 PheGlnSerIleMetAspSerSerGlyGlyLeuGlyLeuProSerAsnLeuAsnThr 112
Db 1220 ACGGCAAGTCTCTCGAGGACTCC-----GGTTTCGGCGACGACTCGGCATC 1267
QY 113 LeuPheSerGlnAlaTrpAsnAla-----GlyAlaArgIleHisThrAsnSerTrpGly 130
Db 1268 CTCGCGGCTGAGTGGGCGCGCGCGAGGGCGGACGTCTCACCATGAGCTCGGGC 1327
QY 131 Ala-----ProValAsnGlyAlaTyrThrAlaAsnSerArgGln 143
Db 1328 GGCATGGACACACCGGAGACCGACCCGCTGGAGCGCGG----- 1366
QY 144 ValAspGluTyrValArgAsnAsnAspMetThrValLeuPheAlaAlaGlyAsnGluGly 163
Db 1367 GTCGACAAGCTGTCGCGCGAGAGGGGTCTCTGTGCCATCGCGCGCGCAACGAGGGC 1426
QY 164 ProAsnSerGlyThrIleSerAlaProGlyThrAlaLysAsnAlaIleThrValGlyAla 183
Db 1427 CCGGAGTGC-----ATCGGTTCGCGCGGAGCGCGGACCGCCCTTCACCTCGCGGCC 1480
QY 184 ThrGluAsnTyrArgProSerPheGlySerIleAlaAspAsnProAsnHisIleAlaGln 203
Db 1481 -----GTCGACGACGAAGCAAGCTCGCGGAC 1507
QY 204 PheSerSerArgGly-----AlaThrArgAspGlyArgIleLysProAspValThrAlaPro 222
Db 1508 TTCTCTCTCCACCGCGCGCGCGCTCGCGGACGCGCCATCAAGCGGAGCGTCCACCGCTCCC 1567
QY 223 GlyThrPheIleLeuSerAlaArgSerSerLeuAlaProAspSerSerPheTrpAlaAsn 242
Db 1568 GCGGTGGACATCACGGCGCGCTCGCGGAGGGCAACGACATCGCGCAGGAGGTGCGTGAG 1627
QY 243 TyrAsnSerLysTyrAlaTyrMetGlyGlyThrSerMetAlaThrProIleValAlaGly 262
Db 1628 GGACCGCGCGGCTACTGACCATCTCGGCGACGTCGATGGCGACCCCGCGCGCTCGCGGGC 1687
QY 263 AsnValAlaGlnLeuArgGluHisPheIleLysAsnArgGlyIleThrProLysProSer 282
Db 1688 GCGCGCGCGCTCTCTGAGCAGCAG-----CACCCCGGAC 1720
QY 283 LeuIleLysAlaAlaLeu-----IleAlaGlyAlaThrAspValGlyLeuGlyTyr 299
Db 1721 TGGACCTCCGCGGAACTGAGGGCGGCTCACCGGCTCCCAAGGGCGGC---AAGTAC 1777
QY 300 ProAsnGlyAspGlnGlyTyrGlyArgValThrLeuAsnLysSerLeuAsnValAlaTyr 319
Db 1778 ACCCGGTTCCGACGAGGTTTCGGCGCGGATCCAGGCCGCAAGGGCGCTCCACGACCGGTG 1837
QY 320 ValAsnGluAlaThrAlaLeuAlaThrGly----- 329
Db 329

QY 378 LysTyrValGlyAsnAspSerTyrProTyrAspAsnAsnTyrAspGlyArgAsnAsn 397
 Db 1471 CAGGTT-----GACTACTCTTAC-----ACCGCTACTATGGA 1503
 QY 398 valGluAsnValPheIleAsnAlaProGlnSerGlyThrTyrIleIleGluValGlnAla 417
 Db 1504 TTCGAAAGGTTGGTTATTACAAACCACTGATGGACATGGACAATTAGGTTGTAAGC 1563
 QY 418 TyrAsnValProSerGlyProGlnArgPheSerLeuAlaIleVal 432
 Db 1564 TAC-----AGCGGAAGTGCACAACTATCATAGTAGATGTGGTA 1599

RESULT 11

US-08-894-818B-6
 ; Sequence 6, Application US/08894818B
 ; Patent No. 6261822

GENERAL INFORMATION:

APPLICANT: TAKAKURA, Hikaru
 APPLICANT: MORISHITA, Mio
 APPLICANT: YAMAMOTO, Katsuhiko
 APPLICANT: MITTA, Masanori
 APPLICANT: ASADA, Kiyozo
 APPLICANT: TSUNASAWA, Susumu
 APPLICANT: KATO, Ikunoshin
 TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Browdy and Neimark
 STREET: 419 Seventh Street N.W., Ste. 300
 CITY: Washington
 STATE: D.C.
 COUNTRY: United States of America
 ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/894,818B
 FILING DATE: 20-MAY-1998

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/JP96/03253
 FILING DATE: 07-NOV-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 323285/1995
 FILING DATE: 12-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Browdy, Roger L.
 REGISTRATION NUMBER: 25,618
 REFERENCE/DOCKET NUMBER: TAKAKURA=1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-5197
 TELEFAX: (202) 737-3528
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1977 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA

US-08-894-818B-6

Alignment Scores:

Pred. No.: 6e-30 Length: 1977
 Score: 395.50 Matches: 129
 Percent Similarity: 44.14% Conservative: 63
 Best Local Similarity: 29.66% Mismatches: 172
 Query Match: 17.50% Indels: 71
 DB: 3 Gaps: 17

US-09-985-689A-4 (1-433) x US-08-894-818B-6 (1-1977)

QY 16 AsnTyrGlyLeuTyrGlyGlnValValAlaValAlaAspThrGlyLeuAspThr 35
 Db 454 AACTTGGGATATGATGTTCTGGAATCAATAAGTAATAATGACATGGAATGAC--- 510
 QY 36 GlyArgAsnAspSerSerMetHisGluAlaPheArgGlyLysIleThrAlaLeu----- 53
 Db 511 -----GCTTCTCATCCAGATCTCCAGGAAAAGTAATGGGTGGGTAGAT 555
 QY 54 TyrAlaLeuGlyArgThrAsnAsnAlaSerAspProAsnGlyHisGlyThrHisValAla 73
 Db 556 TTTGTCAATGCTAGG---AGTTATCCATACGATGACCATGGACATGGAATCATGTAGCT 612
 QY 74 GlySerValLeuGly-----AsnAlaLeuAsnLysGlyMetAlaProGln 88
 Db 613 TCAATAGCAGCTGGTACTGGAGCAGCAAGTAATGGCAAGTACAGGGAATGGCTCCAGGA 672
 QY 89 AlaAsnLeuValPheGlnSerIleMet-----AspSerSerGlyGlyLeuGlyLeu 106
 Db 673 GCTAAGCTGGCGGGAATTAGGTTCTAGTGCCGATGGTTCTGGAAGCATATCTACTATA 732
 QY 107 ProSerAsnLeuAsnThrLeuPheSerGlnAlaTrpAsnAlaGlyAlaArgIleHisThr 126
 Db 733 ATTAAGGGAGTTGAGTGGCGCGTTGATAACAARGATAAGTACGGAATTAAGGTATTAA 792
 QY 127 AsnSerTrpGlyAlaProValAsnGlyAlaTyrThrAlaAsnSerArgGlnValAspGlu 146
 Db 793 CTTTCTCTTGGTTCAGCCAGAGCTCCGACGACCGACTCCCTCAGTCAGCGCGTCAAC 852
 QY 147 TyrValArgAsnAsnAspMetThrValLeuPheAlaAlaGlyAsnGluGlyProAsnSer 166
 Db 853 AACGCTGGGACGCGGTATAGTCTGCGCGCGCAACAGCGGCGCCGCAACACC 912
 QY 167 GlyThrIleSerAlaProGlyThrAlaLysAsnAlaIleThrValGlyValAlaThrGluAsn 186
 Db 913 TACACCGTCGCTCACCCGCGCGAGCAGGTCAATACCGTCGGTGCA----- 963
 QY 187 TyrArgProSerPheGlySerIleAlaAspAsnProAsnHisIleAlaGlnPheSerSer 206
 Db 964 -----GTTGACAGCAACGACACATCGCCAGCTTCTCCAGC 999
 QY 207 ArgGlyAlaThrArgAspGlyArgIleLysProAspValThrAlaProGlyThrPheIle 226
 Db 1000 AGGGGACCGACCGCGGAGGCTCAAGCGGAAGTCTGCGCCCGCGGTGACATC 1059
 QY 227 LeuSerAlaArgSerSerLeuAlaProAspSerPheTrpAlaAsnTyrAsnSerLys 246
 Db 1060 ATAGCCCGCGCGCCAGC-----GGAACCAAGCATGGGCAACCCGATAAACGACTAC 1110
 QY 247 TyrAlaTyrMetGlyGlyThrSerMetAlaThrProIleValAlaGlyAsnValAlaGln 266
 Db 1111 TACACCAAGGCTCTGGAACCAAGCATGCGCCCGCAACCGTTTCGGGCGTTGGCGGCTC 1170
 QY 267 LeuArgGluHisPheIleLysAsnArgGlyIleThrPro-----LysProSerLeu 283
 Db 1171 ATCTTCCAG-----GCCACCCGAGCTGGACCCCGGACAGGTGAAGACCGCCCTC 1221
 QY 284 IleLysAlaAlaLeuIleAlaGlyAlaThrAspValGly-----LeuGlyTyrProAsnGly 302
 Db 1222 ATCGAGACCGCGCAGCATAGTCCGCCCAAGGAGTAGCGGACATCGCCTAC----- 1272
 QY 303 AspGlnGlyTrpGlyArgValThrLeuAsnLysSerLeuAsnValAlaTyrValAsnGlu 322
 Db 1273 -----GGTGGCGGTAGGTGAACGCTTACAGGCC-----ATCAAGTACGACGACTAC 1320
 QY 323 AlaThrAlaLeuAlaThrGlyGln-----LysAlaThrTyrSerPheGln 337
 Db 1321 GCCAAGCTCACCTTCACCGGCTCCGTCGCGCAAGGGAAGCGCCACCCACCTTCGAC 1380
 QY 338 AlaGlnAlaGlyLysProLeuLysIleSerLeuValTrpThrAspAlaProGlySerThr 357
 Db 1381 GTCAGCGCGCCACCTTCTGTGACCCGCCACCTCTACTGGAC----- 1422

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 7.49423 Seconds
(without alignments)
3008.498 Million cell updates/sec

Title: US-09-985-689A-5

Perfect score: 2261
Sequence: 1 NDVARGIVKADVAQNNGLY.....EVQAYNVFSGPQRFSLAIVH 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------|
| 1 | 501 | 22.2 | 1743 | 1 | TAGC DICDI |
| 2 | 463.5 | 20.5 | 1905 | 1 | TAGB DICDI |
| 3 | 312 | 13.8 | 806 | 1 | SUBV BACSU |
| 4 | 307 | 13.6 | 1398 | 1 | PLS PYRFU |
| 5 | 298.5 | 13.2 | 580 | 1 | EXPR_XANCP |
| 6 | 269 | 11.9 | 269 | 1 | PRTM_BACSP |
| 7 | 269 | 11.9 | 269 | 1 | SUBS_BACLE |
| 8 | 269 | 11.9 | 380 | 1 | ELVA_BACAO |
| 9 | 269 | 11.9 | 380 | 1 | ELVA_BACCS |
| 10 | 265 | 11.7 | 534 | 1 | PROA_VIBAL |
| 11 | 263.5 | 11.7 | 420 | 1 | SUBT_BACS9 |
| 12 | 259 | 11.5 | 269 | 1 | SUBS_BACLE |
| 13 | 255 | 11.3 | 401 | 1 | THPS_BACSP |
| 14 | 220 | 11.1 | 894 | 1 | WPRA_BACSU |
| 15 | 244.5 | 10.8 | 1167 | 1 | SCAL_STRPY |
| 16 | 243.5 | 10.8 | 378 | 1 | ELVA_BACSP |
| 17 | 242 | 10.7 | 379 | 1 | SUBT_BACLI |
| 18 | 241.5 | 10.7 | 321 | 1 | ISP_BACCS |
| 19 | 241 | 10.7 | 513 | 1 | AQLI_THAQ |
| 20 | 237.5 | 10.5 | 422 | 1 | TKSU_PVRKO |
| 21 | 237 | 10.5 | 1181 | 1 | SCA2_STRPY |
| 22 | 234.5 | 10.4 | 402 | 1 | ALP_CEPAC |
| 23 | 231.5 | 10.2 | 1052 | 1 | MSIP_HUMAN |
| 24 | 228.5 | 10.1 | 319 | 1 | ISP1_BACSU |
| 25 | 228 | 10.1 | 1052 | 1 | MSIP_CRIGR |
| 26 | 228 | 10.1 | 1052 | 1 | MSIP_MOUSE |
| 27 | 228 | 10.1 | 1052 | 1 | MSIP_RAT |
| 28 | 227.5 | 10.1 | 381 | 1 | SUBN_BACNA |
| 29 | 227.5 | 10.1 | 381 | 1 | SUBT_BACSA |
| 30 | 227.5 | 10.1 | 381 | 1 | SUBT_BACST |
| 31 | 226.5 | 10.0 | 381 | 1 | SUBT_BACSU |
| 32 | 225.5 | 10.0 | 1433 | 1 | SUBF_BACSU |
| 33 | 225 | 10.0 | 382 | 1 | SUBT_BACAM |

34 224.5 9.9 275 1 SUBT_BACPU
35 224 9.9 274 1 SUBD_BACLI
36 223.5 9.9 279 1 THET_THEVU
37 223 9.9 595 1 BPRX_BACNO
38 220 9.7 404 1 SMP1_MAGPO
39 219 9.7 388 1 CUDP_METAN
40 217.5 9.6 326 1 ISP_PAEPO
41 213.5 9.4 409 1 ALP_TRIHA
42 212 9.4 533 1 PEPC_ASPNG
43 211 9.3 451 1 ISPE_SCHPO
44 209.5 9.3 451 1 XLP1_SCHPO
45 209 9.2 530 1 HLY_HALI7

ALIGNMENTS

RESULT 1

TAGC DICDI STANDARD; PRT; 1743 AA.
AC Q23868;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Prestalk-specific protein tagC precursor (EC 3.4.21.-).
GN TAGC.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=97140317; PubMed=8986798;
RA Shaulsky G., Escalante R., Loomis W.F.;
RT "Developmental signal transduction pathways uncovered by genetic
suppressors.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:15260-15265(1996).
CC -!- FUNCTION: Intercellular communication via tagC may mediate
integration of cellular differentiation with morphogenesis (By
similarity).
CC -!- SIMILARITY: In the N-terminal section; belongs to peptidase family
S8.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING
TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -!- SIMILARITY: STRONG, TO TAGB.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U60086; AAB03331.1; -.
CC PIR; T18279; T18279.
CC DictyBase; DDB0001795; tagC.
CC InterPro; IPR003593; AAA ATPase.
CC InterPro; IPR001140; ABC_TM_transp.
CC InterPro; IPR003439; ABC_transporter.
CC InterPro; IPR000209; Peptidase_S8.
CC Pfam; PF00664; ABC_membrane; 1.
CC Pfam; PF00005; ABC_tran; 1.
CC Pfam; PF00082; Peptidase_S8; 1.
CC PRINTS; PR00723; SUBTILISIN.
CC SMART; SMO0382; AAA; 1.
CC PROSITE; PS00929; ABC_TM1P; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_2; 1.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
CC PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
CC PROSITE; PS00137; SUBTILASE_HIS; 1.
CC PROSITE; PS00138; SUBTILASE_SER; FALSE_NEG.
CC Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane;
KW

FT SIGNAL 1 31
FT CHAIN 32 1905
FT DOMAIN 378 700
FT DOMAIN 1518 1756
FT TRANSMEM 1011 1031
FT TRANSMEM 1076 1096
FT TRANSMEM 1121 1141
FT TRANSMEM 1210 1230
FT TRANSMEM 1309 1329
FT TRANSMEM 1332 1352
FT ACT_SITE 387
FT ACT_SITE 432
FT ACT_SITE 695
FT NP_BIND 1553
FT DOMAIN 63 67
FT DOMAIN 95 104
FT DOMAIN 107 134
FT DOMAIN 311 321
FT DOMAIN 833 837
FT DOMAIN 838 844
FT DOMAIN 871 876
FT DOMAIN 1012 1015
FT DOMAIN 1386 1389
FT DOMAIN 1398 1404
FT DOMAIN 1445 1450
FT DOMAIN 1765 1779
FT DOMAIN 1782 1785
FT DOMAIN 1807 1812
FT DOMAIN 1813 1860
FT DOMAIN 1872 1878
FT CARBOHYD 594 594
FT CARBOHYD 621 621
FT CARBOHYD 672 672
FT CARBOHYD 747 747
FT CARBOHYD 823 823
FT CARBOHYD 1172 1172
FT CARBOHYD 1522 1522
FT CARBOHYD 1658 1658
SQ SEQUENCE 1905 AA; 212518 MW; B8E223FA8B9AE13C CRC64;

Query Match
Best Local Similarity 20.5%; Score 463.5; DB 1; Length 1905;
Matches 162; Conservative 78; Mismatches 155; Indels 183; Gaps 28;

QY 19 LYGGQGVAVATGIDTGR---NDS-----SMHEAFRGKITALYALGRNTNANDP--N 66
DB 376 LGKGQILSIATGIDGCHCFSDSKYPIPFQVNVNHRKVYI-----TYHDNEDYV 430
QY 67 GHGTHVAGSVLGN-----ALN--KGMAPQANLVFQSIQNDSSGGLGL-PSNLNTLFQ 117
DB 431 GHGTHVCGSAAGTPEDSSWAISFSGLATDAXIAFYDL--SSGSSEPTPPEDYSQMKPL 488
QY 118 WNAARIHTNSGA---PVNGAYTANSQVDEYV-RNNDMTVLFAAGNEGPNSTISAP 172
DB 489 YPAGARVHGDSVGSVLSQYGYGSDAGGIDAFLEYEPFSLRAAGN-NELFASLLAQ 547
QY 173 GTAKNAITYGATE---NY-----RP----- 189
DB 548 ATAKNAITYGAEQTAHVNVVSDALEYDFSDNANFQPCLPDKKYCNVTTAKCCSEVNV 607
QY 190 -----SFGSLAD--NPNHIAQFSRGNTRGRIPKDPVTPAGTILSAR 230
DB 608 KGLQLCCPASIKQNASDSTTQPFYNNENMNGSFSGKGTGRLKPDIVAPGEYITSAR 667
QY 231 SS-----LAPDSSFWANTNSKYAVMGTSMTPTVAGNVQAQLRHF-----IKN 274
DB 668 SNGENSTDCGDSGL-PNANGLMS-ISGTSMTPLATAATTILRQYLVGDYPTGESVEE 725
QY 275 RGITPKPSIKKALINAGATVGLGY-----PSGD-----CGWGRVTLDKSL 315
DB 726 NKLLPTGSLIKALMINNAQLNGTYFWSASTNPSNAIFEQINGANLIQGWGALEMMNWL 785
QY 316 NVAYVN-----EAT-----ALATGQKATYSFQAQ-----AGK 342

DB 786 YVKSSNPTPPSRWIGIGGLGKQKATFEMKEDSLSSGLNKSCTYKTPSSSSSGSGGGGT 845
QY 343 P-LKISLVMTDAPGTTASTYTIWDLDL-----VITAPN--GQKVGNDFSYPDN 390
DB 846 PRIVATLVMTDPPSYSGAKFNLVNLDLLLSDDDDSIITIGNSGSLQPAKVAQP--- 902
QY 391 NMDGRNNVENVFNAPOSGTYTIEVQAVNVPSGPQRF 428
DB 903 --DTLNNVEGIILNPTKANNYKFTIAGTNVPIGPQKFS 938
RESULT 3
SUBV_BACSU
ID SUBV_BACSU STANDARD; PRT; 806 AA.
AC P29141;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Minor extracellular protease vpr precursor (EC 3.4.21.-).
GN VPR OR IPA-45R OR BSU38090.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 161-195.
RX MEDLINE=92041574; PubMed=1938892;
RA Sloma A., Rufo G.A. Jr., Theriault K.A., Dwyer M., Wilson S.W.,
RA Pero J.;
RT "Cloning and characterization of the gene for an additional
RT extracellular serine protease of Bacillus subtilis.";
RL J. Bacteriol. 173:6889-6895 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=95020537; PubMed=7934828;
RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,
RA Hullo M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,
RA Presecan E., Santana M., Schneider E., Schweizer J., Vertes A.,
RA Rapoport G., Danchin A.;
RA "Bacillus subtilis genome project: cloning and sequencing of the 97
RT kb region from 325 degrees to 333 degrees.";
RL Mol. Microbiol. 10:371-384 (1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Avevedo V., Betero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Etian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maestl C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noore D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Pressecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Taccini E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzensegger T.,

```

Db      519 HDPDPHY--GYGSKQ-----GTSMASPHIAGAVAIKQ-----AKPKWSVEQIKAAIMN 565
Qy      291 GATDV----GLGYPSGQGGWRVTLDKSLNVAVNEATLATGQKATY-SFOAQAGKPLK 345
Db      566 TAVTLKDSGDGVYVPHNAQAG-----SARIMNAIKADSLVSPGYSYGTFLKENGNETK 619
Qy      346 ISLVMTDAFGSTTASYTLVNDLDLVTAPNGQKVGNDSPSYVDNWDNRNVNVEFNA 405
Db      620 NETFTIENQSGIRKSYTL-----EYFNGSGISTGTSRVVIPA 658
Qy      406 PQSGYTTIEVQ 416
Db      659 HQTGKATKVK 669

RESULT 4
IDS_PYRPU
PLS_PYRPU STANDARD; PRT; 1398 AA.
AC P72186;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pyrolysin precursor (EC 3.4.21.-).
GN PLS OR PF0287.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OC NCBI_TaxId=2261;
[1]
SEQUENCE FROM N.A., SEQUENCE OF 150-184, AND CHARACTERIZATION.
RP STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RX MEDLINE=96355370; PubMed=8702780;
RA Voorthorst W.G.B., Edgen R.I.L., Geerling A.C.M., Platteeuw C.,
RA Sierzen R.J., de Vos W.M.
RT "Isolation and characterization of the hyperthermostable serine
RT protease, pyrolysin, and its gene from the hyperthermophilic archaeon
RT Pyrococcus furiosus.";
RL J. Biol. Chem. 271:20426-20431(1996).
[2]
SEQUENCE FROM N.A.
RP STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RX STRAIN=Vc1, Dunn D.M., Robb F.T., Brown J.R.;
RA Weiss R.B., Voorthorst W.G.B., Dijkgraaf M., Kluskens L.D.,
RL "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
[3]
CHARACTERIZATION, AND 3D-STRUCTURE MODELING.
RP MEDLINE=21079021; PubMed=11210516;
RX de Vos W.M., Voorthorst W.G.B., Dijkgraaf M., Kluskens L.D.,
RA Van der Oost J., Siezen R.J.;
RT "Purification, characterization, and molecular modeling of pyrolysin
RT and other extracellular thermostable serine proteases from
RT hyperthermophilic microorganisms.";
RL Meth. Enzymol. 330:383-393(2001).
CC -1- FUNCTION: Has endopeptidase activity toward caseins, casein
CC fragments including alpha-S1-casein and synthetic peptides.
CC -1- SUBCELLULAR LOCATION: Cell envelope associated.
CC -1- PTM: LMW pyrolysin seems to be produced by autoproteolytic
CC activation of HMW pyrolysin.
CC -1- PTM: Glycosylated.
CC -1- MISCELLANEOUS: Thermostable; high activity at 95 degrees Celsius.
CC -1- SIMILARITY: Belongs to peptidase family S8.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U55835; AA09761.1; -.
DR EMBL; AR010153; AAL80411.1; -.
PIR; T28159; T28159

```

DR HSP; Q45670; IDBI.
 DR MEROPS; S08.100; -
 DR InterPro; IPR000203; Peptidase_S8.
 DR InterPro; IPR007280; PPC.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE SER; 1.
 KW Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
 KW Complete proteome.
 FT SIGNAL 1 26 POTENTIAL.
 FT PROPEP 27 149
 FT CHAIN 150 1398
 FT ACT_SITE 179 179
 FT ACT_SITE 365 365
 FT ACT_SITE 590 590
 FT CARBOHYD 152 152
 FT CARBOHYD 222 222
 FT CARBOHYD 228 228
 FT CARBOHYD 240 240
 FT CARBOHYD 257 257
 FT CARBOHYD 262 262
 FT CARBOHYD 298 298
 FT CARBOHYD 327 327
 FT CARBOHYD 406 406
 FT CARBOHYD 651 651
 FT CARBOHYD 663 663
 FT CARBOHYD 739 739
 FT CARBOHYD 792 792
 FT CARBOHYD 893 893
 FT CARBOHYD 908 908
 FT CARBOHYD 917 917
 FT CARBOHYD 929 929
 FT CARBOHYD 1048 1048
 FT CARBOHYD 1056 1056
 FT CARBOHYD 1084 1084
 FT CARBOHYD 1117 1117
 FT CARBOHYD 1133 1133
 FT CARBOHYD 1140 1140
 FT CARBOHYD 1148 1148
 FT CARBOHYD 1208 1208
 FT CARBOHYD 1233 1233
 FT CARBOHYD 1237 1237
 FT CARBOHYD 1332 1332
 FT CONFLICT 607 609
 FT CONFLICT 881 881
 FT SEQUENCE 1398 AA; 154474 MW; 355D873A27D56552 CRC64;
 Query Match 13.6%; Score 307; DB 1; Length 1398;
 Best Local Similarity 28.4%; Pred. No. 2.9e-12;
 Matches 129; Conservative 37; Mismatches 153; Indels 136; Gaps 15;
 21 GGGVVAVADTGLDGRNDS-----SMHAFRGKITALYALGRNNANDPN-- 66
 301 GNGYDIAYVDTLDVDFDEPLGQVNTVYDAVFYSYGLN--YVLAEL----DPNGE 354
 67 -----GCHGTHVAGSVLGNALN-----KG 84
 355 YAVFGDGHGTHVAGTVAGVDSNDNDWLSMYSGEWEVSRYLGDWYTNVTTDTVQG 414
 85 MAPQANLVFQSIINDSGGIGGLPSNLNTLFSQANAGARIHTNSWG--APVNGAYTANSR 142
 415 VAPGAQIWAIRVLRSDG--RGSMDWDIEGMTVAATHGADVISMSLGGNAPVLDGTDPSV 472
 143 QVDEYVRNNDMTVLFAAGNEGNSGTISAPGAKNAITVGATE----- 185
 473 ANDELTEKGVVFIAGNEGPGCINIVGSPGVATKAITVGAAPVFNVGYYVQALGYPD 532
 186 ----NYRPSFGSLADPNHIAFSSRGATDRGIKPDVTPAGTFTILSRSLAPDSFWA 241
 533 YYGFFYFPATVNV----RIAFSSRGPRIDGEIKPNVAPGYGIYSILPMWIGADF-- 585

QY 242 NYSKYAYMGTSWATPIVAGNVAQLREHFKRGITPKPSLIKAALIAGAT-----DV 295
 Db 586 -----MSGTSWATPHVSGWALLISG-AKAGIYYNPDIITKKVLESATWLESGDPYT 636
 QY 296 GLGYPSGPDQGWRTLDKSLNVVNEATATATCQKATYSFOAQAGKFLKISLVTWTDAPG 355
 Db 637 GQKYTELDQGHGLVNVTKSWEI-----LKAINGTTLTIVDHWADKSY 678
 QY 356 STTASYTLVNDLDLVIITAPNG-----OKYGVN 382
 Db 679 SDFAEYLGVDVIRGLYARNIPDIVWEHIKYVD 712

RESULT 5
 EXPR_XANCP
 ID_EXPR_XANCP STANDARD; PRT; 580 AA.
 AC F23314;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, last sequence update)
 DT 28-FEB-2003 (Rel. 41, last annotation update)
 DE Extracellular protease precursor (EC 3.4.21.-).
 GN XCC0851.
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 EX MEDLINE=90251253; PubMed=2187155;
 RA Liu Y.-N., Tang J.-L., Clarke B.R., Dow J.M., Daniels M.J.;
 RT "A multipurpose broad host range cloning vector and its use to
 RT characterise an extracellular protease gene of Xanthomonas campestris
 RT pathovar campestris.";
 RL Mol. Gen. Genet. 220:433-440(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Cimarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.T.T.,
 RA Fomighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi I.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities.";
 RL Nature 417:459-463(2002).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; X51635; CAA35962.1; -.
 DR EMBL; AE012184; AAM40166.1; -.
 DR PIR; S11890; S11890.
 DR HSP; P00782; 2S8T.
 DR MEROPS; S08.UFA; -.

```

DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR InterPro; IPR009020; Protease_inhib.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR Hydrolyase; Serine protease; Zymogen; signal; Complete proteome.
FT SIGNAL 1 32 POTENTIAL.
FT PROPEP 33 2136 POTENTIAL.
FT CHAIN 2137 580 EXTRACELLULAR PROTEASE.
FT ACT_SITE 177 177 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 237 237 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 409 409 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 225 273 BY SIMILARITY.
FT DISULFID 315 352 BY SIMILARITY.
FT DISULFID 450 454 POTENTIAL.
SQ SEQUENCE 580 AA; 57228 MW; 8C9A2CAE4B7F47CB CRC64;

Query Match 13.2%; Score 298.5; DB 1; Length 580;
Best Local Similarity 29.0%; Pred. No. 3.3e-12;
Matches 137; Conservative 51; Mismatches 149; Indels 135; Gaps 26;

QY 21 GQSQVAVADTGL-----D*GRNDSMHEAFRGKITALYALGRITNAND----- 64
DB 168 GSGTVVAVIDTGTITSHADLNANLAGYDFISATTDARDGNGRDSNADEGDHYAANECA 227

QY 65 -----PNGHGHVAGS-----VLGNALNKGMAPOANL-----VFOSIMDS 99
DB 228 GIPAASSWHGTHVAGTAAVTNNTTGVAGTAYGAKVFPVRLGKGGSLSDIADAIYWA 287

QY 100 SGG-LGGLPNLN--TLFSDAMNAGARIHTNSGAPVNGAYTANSRQVDVYVRNNDMTVL 156
DB 288 SGTVSGIPANAPAEVINNSLGGGSCSTTWNA--INGAVSRGT-----TVV 334

QY 157 FAAGNEGPN--SGTISAPGTAKNAITVGATENYRPSGLADNPNHTAQPSRGATRDGRI 215
DB 335 VAAGNDASNVSG--SLPANCANVIAVAAT-----TSAGAK-----ASYSNFGT----- 375

QY 216 KPDVTAPGTPLSARSS--LAPDSFMANVSKYAYMGTSNATPIVAGNVQAQLREHFIK 273
DB 376 GIDVSAFPGSILTLNSGTTTPGASAYSN-----GTSNAPSHVAGVALVQS--VA 426

QY 274 NRGITPK--PSLIK--AALIAGATDVGLGY-----PSGDQGWGRVTLDKS 314
DB 427 PTALTPAAVETLLKNTARALPGACSGCGAGIVNADAATAVTAALNGSGGGGGGNTLTNG 486

QY 315 LNVAYVNEATALATGQATYSFQAQKPKLSLWTDAPGSTTASYTLVNDLVLV---- 371
DB 487 TPVT-----GLGAATGAELNNTIITVPAGSG---TLVTTSGGG-----GDADLYVRAG 531

QY 372 TAPNGQKYGVNDFSPYDNNWDRNVENVFINAPQSFTYTIIEVQAYNVPSG 423
DB 532 SAPTDSAYT---CRPVRS-----GNAETCTITAP-SGYTYVRLKAYSTFSG 573

RESULT 6
PRTM_BACSP STANDARD; PRY; 269 AA.
AC Q99405;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE M-protease (EC 3.4.21.-).
OS Bacillus sp. (strain KSM-K16).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RA Yanane T., Kani T., Hatanaka T., Suzuki A., Ashida T.,
RA Kobayashi T., Ito S., Yamashita O.;

```

```

FT "Structure of a new alkaline serine protease (M-protease) from
RT Bacillus sp. KSM-K16."
RL Acta Crystallogr. D 51:199-206(1995).
RN [2]
RX MEDLINE=95358832; PubMed=7632397;
RA Kobayashi T., Hakamada Y., Adachi S., Hitomi J., Yoshimatsu T.,
RA Koike K., Kawai S., Ito S.;
RT "Purification and properties of an alkaline protease from
RT alkalophilic Bacillus sp. KSM-K16."
RL Appl. Microbiol. Biotechnol. 43:473-481(1995).
CC -1- COPACTOR: Binds 2 calcium ions per subunit.
CC -1- SUBUNIT: Monomer.
CC -1- SIMILARITY: Belongs to peptidase family S8.
DR MEROPS; S08.010; -.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Serine protease; Metal-binding; Calcium-binding;
KW 3D-structure.
FT ACT_SITE 32 32 CHARGE RELAY SYSTEM.
FT ACT_SITE 62 62 CHARGE RELAY SYSTEM.
FT ACT_SITE 215 215 CHARGE RELAY SYSTEM.
FT METAL 2 2 CALCIUM 1.
FT METAL 40 40 CALCIUM 1.
FT METAL 73 73 CALCIUM 1.
FT METAL 75 75 CALCIUM 1.
FT METAL 77 77 CALCIUM 1.
FT METAL 79 79 CALCIUM 1.
FT METAL 163 163 CALCIUM 1 (VIA CARBONYL OXYGEN).
FT METAL 165 165 CALCIUM 1 (VIA CARBONYL OXYGEN).
FT METAL 168 168 CALCIUM 1 (VIA CARBONYL OXYGEN).
FT METAL 168 168 CALCIUM 2 (VIA CARBONYL OXYGEN).
FT METAL 168 168 CALCIUM 2 (VIA CARBONYL OXYGEN).
FT HELIX 6 10
FT TURN 11 12
FT HELIX 13 18
FT TURN 19 20
FT TURN 24 25
FT TURN 27 32
FT TURN 39 40
FT STRAND 43 48
FT TURN 51 52
FT HELIX 62 71
FT TURN 84 85
FT STRAND 87 92
FT TURN 96 97
FT HELIX 102 114
FT TURN 115 116
FT STRAND 119 122
FT STRAND 126 126
FT HELIX 131 142
FT TURN 143 144
FT STRAND 146 150
FT STRAND 161 161
FT TURN 162 164
FT TURN 166 167
FT STRAND 169 174
FT TURN 176 177
FT STRAND 180 180
FT TURN 182 183
FT TURN 188 189
FT STRAND 192 195
FT STRAND 199 203
FT TURN 204 206
FT STRAND 207 211
FT HELIX 214 231
FT TURN 233 234
FT TURN 237 246
FT TURN 247 247
FT STRAND 249 249

```

```

FT 254 257
FT TURN 258 259
FT STRAND 261 261
FT HELIX 264 267
FT TURN 268 269
SQ SEQUENCE 269 AA; 26723 MW; 7A03C86D534A1D07 CRC64;

Query Match 11.9%; Score 269; DB 1; Length 269;
Best Local Similarity 31.7%; Pred.No.1e-10;
Matches 98; Conservative 32; Mismatches 97; Indels 82; Gaps 14

13 AONNYGLYGQGVAVADTGLDTCRNDSSMHEAFRGKITALYALGRITNNANDPENGHGTHV 72
15 AAHNEGLTSGVKVAVLDTGTST-----HPDLNIRGGASFPVGP-STDONGHGTHV 66
73 AGSV--LGNALN-KGMAPQANLVFQSIWDSGGGLGSPNLNTLFSQAWNAGARIHTNSW 129
67 ACTIAALNNSIGVLGAPSAELYAVKVLGASG--SGSVSIIAQGLEWAGNGHVAANLSL 124
130 GAPVNGAGTANSRQVDYEVRRNDMTVLFAAGNEGPNSTGISAPGTAKNALTYGATENYEP 189
125 GSP----SFSATLEQVNSGATRGVIVVAASGNSG--AGSISYFARYANAWAVCAT----- 174
190 SFGSLADPNPHIAFSSRGATRGDKRIKFDVTAPOGTILTSARSSLAPDSSFWMYANSKYAY 249
175 -----DQNNRASPQYAGL-----DIVAPGVNQSTYP-----GSTYAS 210
250 MGGTSWATPIVAGNVA-----QUREHEFIKRGITPKPSLIKAAIAGATDVGL 297
211 LNGTSMATPHVAGVAALYKQKPSNSVQVNRNH-LRN-----TATGL 251
298 G----YFSG 302
252 GNTNLYSG 260

RESULT 7
SUBS BACLE
ID SUBS BACLE STANDARD; PRT; 269 AA.
AC P29600;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Subtilisin Savinase (EC 3.4.21.62) (Alkaline protease).
OS Bacillus lentus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxID=1467;
RN [1]_TaxID=1467;
RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).
RX MEDLINE=92148829; PubMed=1738156;
RA Betzel C., Klupsch S., Papendorf G., Hastrup S., Branner S.,
RA Wilson K.S.;
RT "Crystal structure of the alkaline proteinase Savinase from Bacillus
RT lentus at 1.4-A resolution.";
RL J. Mol. Biol. 223:427-445(1992).
RN [2]
RP STRUCTURE BY NMR.
RX MEDLINE=96184541; PubMed=8654411;
RA Remerowski M.L., Pepermans H.A.M., Hilbers C.W., van de Ven F.J.M.;
RT "Backbone dynamics of the 269-residue protease Savinase determined
RT from 15N-NMR relaxation measurements.";
RL Eur. J. Biochem. 235:629-640(1996).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (0.78 ANGSTROMS).
RX MEDLINE=98426039; PubMed=9753430;
RA Kuhn P., Knapp M., Soltis S.M., Ganshaw G., Thoene M., Bott R.;
RT "The 0.78-A structure of a serine protease: Bacillus lentus
RT subtilisin.";
RL Biochemistry 37:13446-13452(1998).
CC -!- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
CC it catalyzes the hydrolysis of proteins and peptide amides.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large uncharged residue
```

```

FT STRAND 207 211
FT HELIX 214 231
FT TURN 233 234
FT HELIX 237 246
FT TURN 247 247
FT STRAND 249 249
FT HELIX 254 257
FT TURN 258 259
FT STRAND 261 261
FT HELIX 264 267
FT TURN 268 269
SQ SEQUENCE 269 AA; 26698 MW; 4D89F877899BF8D CRC64;

Query Match 11.9%; Score 269; DB 1; Length 269;
Best Local Similarity 31.8%; Pred. No. 1e-10; Indels 74; Gaps 14;
Matches 97; Conservative 35; Mismatches 99;

QY 13 AQNNYGLYGGQVAVAVADTGLDTRNDSSMHEAFKRGKITALYALGRTNNANDPNHGHTHV 72
Db 15 AAHNRGLTSGVGVAVLDTGIST-----HPDLNIRGCASFVGPGE-STDGNGHGHTHV 66

QY 73 AGSV--LGNALN-KGVAPQANLVFQSIMDSSGGLGPSNLNTLFSQAWNAGARITHNSW 129
Db 67 AGTIALNNSIGVLGVAPSAELYAVKVLGASG--SGSVSSIAQGLEWAGNNGMHVANLSL 124

QY 130 GAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTGISAPGTAKNAITVGATENYRP 189
Db 125 GSP---SPSATLFCQAVNSATSRGVLVVAASGNSG--AGSISYPARYANAMAVGAT----- 174

QY 190 SFGSLADPNHIAQFSRGATRDGRKPKDVPATGTFILARSGLAPDSSFANYSKAY 249
Db 175 -----DQNNRASFSQYGAGL-----DIVAPGVNVQSTYP-----GSTVAS 210

QY 250 MGGTSMATPTVAGNVA-----OLRHFKNRGITPKPSLIKAALTAGATDVL 297
Db 211 LNTGSMATPHVAGAAALVKQKNSWNVQIRNH-LKN-----TATSLGSLNL-- 256

QY 298 GYPGG 302
Db 257 -YGGG 260

RESULT 8
ELYA_BACAO STANDARD; PRT; 380 AA.
AC P27693;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alkaline protease precursor [EC 3.4.21.-].
OS Bacillus alcalophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1445;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=PB92;
EX MEDLINE=91282483; PubMed=2059048;
RA van der Leen J.C., Gerritse G., Mulleners L.J.M., van der Hoek R.A.,
Quax W.J.;
FT "Cloning, characterization, and multiple chromosomal integration of a
RT Bacillus alkaline protease gene.";
RL Appl. Environ. Microbiol. 57:901-909(1991).
RN [2]
RC X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
RC STRAIN=PB92;
RX MEDLINE=92390330; PubMed=1518788;
RA van der Leen J.C., Tepiyakov A.V., Kelders H., Kalk K.H., Misset O.,
RA Mulleners L.J.M., Dijkstra B.W.;
RT "Crystal structure of the high-alkaline serine protease PB92 from
RT Bacillus alcalophilus.";
RL Protein Eng. 5:405-411(1992).
RN [3]
RN X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).

```

```

RX MEDLINE=93078250; PubMed=1447775; Schomburg D.;
RA Sobek H., Hecht H.-J., Ahle W., Schomburg D.;
RT "X-ray structure determination and comparison of two crystal forms of
RT a variant (Asn115Arg) of the alkaline protease from Bacillus
RL alcalophilus refined at 1.85-A resolution.";
RN J. Mol. Biol. 228:108-117(1992).
RP [4]
RP STRUCTURE BY NMR OF 112-380.
RC STRAIN=PB92;
RX MEDLINE=97272737; PubMed=9115441;
RA Martin J.R., Mulder F.A., Karimi-Nejad Y., van der Zwan J.,
RA Mariani M., Schipper D., Boelens R.;
RT "The solution structure of serine protease PB92 from Bacillus
RT alcalophilus presents a rigid fold with a flexible substrate-binding
RT site.";
RL Structure 5:521-532(1997).
CC -!- COFACTOR: Binds 2 calcium ions per subunit.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to peptidase family S8.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M65086; AAA22212.1; -.
DR EMBL; A13738; CAA01128.1; -.
DR PIR; A49778; A49778.
DR PDB; 1AH2; 15-APR-98.
DR MEROPS; S08.038; -.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR009020; Protease_inhib.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PS00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
KW Signal; 3D-structure.
FT SIGNAL 1 27
FT PROPEP 28 112
FT CHAIN 113 380
FT ACT_SITE 143 143
FT ACT_SITE 173 173
FT ACT_SITE 326 326
FT METAL 113 113
FT METAL 151 151
FT METAL 184 184
FT METAL 186 186
FT METAL 188 188
FT METAL 190 190
FT METAL 274 274
FT METAL 276 276
FT METAL 279 279
FT METAL 117 120
FT TURN 121 123
FT HELIX 125 130
FT TURN 131 131
FT TURN 135 136
FT STRAND 138 142
FT TURN 150 151
FT STRAND 156 157
FT HELIX 173 182
FT STRAND 198 201
FT TURN 207 208
FT HELIX 213 226
FT STRAND 230 233
FT HELIX 242 252
FT TURN 253 255
FT STRAND 257 258

```

```

ALKALINE PROTEASE.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CALCIUM 1.
CALCIUM 1.
CALCIUM 1 (VIA CARBONYL OXYGEN).
CALCIUM 1.
CALCIUM 1 (VIA CARBONYL OXYGEN).
CALCIUM 1 (VIA CARBONYL OXYGEN).
CALCIUM 2 (VIA CARBONYL OXYGEN).
CALCIUM 2 (VIA CARBONYL OXYGEN).
CALCIUM 2 (VIA CARBONYL OXYGEN).
POTENTIAL.

```

FT TURN 267 268
 FT TURN 272 275
 FT STRAND 279 279
 FT STRAND 282 285
 FT STRAND 291 291
 FT TURN 299 300
 FT STRAND 303 306
 FT STRAND 310 314
 FT TURN 315 317
 FT STRAND 318 322
 FT STRAND 325 342
 FT TURN 344 345
 FT HELIX 348 358
 FT HELIX 375 380
 SQ SEQUENCE 380 AA; 38853 MW; 539EA72771B6682C CRC64;

Query Match 11.9%; Score 269; DB 1; Length 380;
 Best Local Similarity 31.8%; Pred. No. 1.6e-10;
 Matches 97; Conservative 35; Mismatches 99; Indels 74; Gaps 14;

QY 13 AQNNYGLYGGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRITNNANDPNGHGHV 72
 Db 126 AAHNRGLTSGVGVAVLDTGIST-----HPDLNIRGGASFVGP-STDQNGHGHV 177

QY 73 AGSV--LGNALN-KGMAPQANLVFQSIMDSSGGLGLPSNLTLFSQAWNAGARIHTNSW 129
 Db 178 AGTAAALNSIGVLGVAPNAELVAVKVLGASG--SGSVSSIAQGLEWAGNGMHVANLSL 235

QY 130 GAPVNGAYTANSRQVDEVRNDMTVLFAGNEGPNSTISAPGTAKNAITVGATENYRP 189
 Db 236 GSP---SPSATLEQAVNSATSRGVLVAAASNGS--AGSISYPARYANAMAVGAT----- 285

QY 190 SPGLADNPNIHIAQFSRGATDRGRIKPDVTAPGTFFILSARSSLAPDSSFWMYNSKYAY 249
 Db 286 -----DQNNRASFQYAGL-----DIVAPGVNVQSTYP-----GSTVAS 321

QY 250 MGGTSMATPIVAGNVA-----QLREHFIKNGRITPKPSLIKAALIAGATDVL 297
 Db 322 LNGTSMATPHVAGAAALVKQNPFSWNSVQIRNH-LKN-----TATSLGSLNLT--- 367

QY 298 GYPSG 302
 Db 368 -YGSG 371

RESULT 9

ELIYA_BACCS STANDARD; PRT; 380 AA.
 ID ELIYA_BACCS AC P41362;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Alkaline protease precursor (EC 3.4.21.-).
 OS Bacillus clausii.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=79880;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=221 / ATCC 21522 / JCM 9139 / DSM 2512;
 RX MEDLINE=93043753; PubMed=1368952;
 RA Takami H., Kobayashi T., Kobayashi M., Yamamoto M., Nakamura S.,
 RA Aono R., Horikoshi K.;
 RA "Molecular cloning, nucleotide sequence, and expression of the
 RT structural gene for alkaline serine protease from alkaliphilic
 RT Bacillus sp. 221.";
 RL Biosci. Biotechnol. Biochem. 56:1455-1460(1992).
 RN [2]
 RP SEQUENCE OF 112-129.
 RC STRAIN=221 / ATCC 21522 / JCM 9139 / DSM 2512;
 RA Horikoshi K.;
 RA (In) Horikoshi K. (eds.);
 RL Microorganisms in alkaline environments, pp.187-194, VCH,
 RL Weinheim (1991).

CC -1- COFACTOR: Binds 2 calcium ions per subunit (By similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; S48754; AAC60420.1; -;
 DR EMBL; D13157; BAA02442.1; -;
 DR EMBL; A26817; CAA01836.1; -;
 DR EMBL; A22550; CAA01611.1; -;
 DR HSSP; P29600; 1GCI.
 DR MEROPS; S08.103; -;
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR009020; Protease_inhib.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PRO0723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Hydrolase; Serine protease; 2; ymogen; Metal-binding; Calcium-binding;
 KW Signal.
 KW SIGNAL 1 27 POTENTIAL.
 FT PROPEP 28 111 ALKALINE PROTEASE.
 FT CHAIN 112 380 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 143 143 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 173 173 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 326 326 CALCIUM 1 (BY SIMILARITY).
 FT METAL 113 113 CALCIUM 1 (BY SIMILARITY).
 FT METAL 151 151 CALCIUM 1 (VIA CARBONYL OXYGEN)
 FT METAL 184 184 (BY SIMILARITY).
 FT METAL 186 186 CALCIUM 1 (BY SIMILARITY).
 FT METAL 188 188 CALCIUM 1 (VIA CARBONYL OXYGEN)
 FT METAL 190 190 (BY SIMILARITY).
 FT METAL 274 274 CALCIUM 1 (VIA CARBONYL OXYGEN)
 FT METAL 276 276 (BY SIMILARITY).
 FT METAL 276 276 CALCIUM 2 (VIA CARBONYL OXYGEN)
 FT METAL 279 279 (BY SIMILARITY).
 FT METAL 279 279 CALCIUM 2 (VIA CARBONYL OXYGEN)
 FT SEQUENCE 380 AA; 38826 MW; 5F73ABC68D5B6831 CRC64;

Query Match 11.9%; Score 269; DB 1; Length 380;
 Best Local Similarity 31.8%; Pred. No. 1.6e-10;
 Matches 97; Conservative 35; Mismatches 99; Indels 74; Gaps 14;

QY 13 AQNNYGLYGGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRITNNANDPNGHGHV 72
 Db 126 AAHNRGLTSGVGVAVLDTGIST-----HPDLNIRGGASFVGP-STDQNGHGHV 177

QY 73 AGSV--LGNALN-KGMAPQANLVFQSIMDSSGGLGLPSNLTLFSQAWNAGARIHTNSW 129
 Db 178 AGTAAALNSIGVLGVAPNAELVAVKVLGASG--SGSVSSIAQGLEWAGNGMHVANLSL 235

QY 130 GAPVNGAYTANSRQVDEVRNDMTVLFAGNEGPNSTISAPGTAKNAITVGATENYRP 189
 Db 236 GSP---SPSATLEQAVNSATSRGVLVAAASNGS--AGSISYPARYANAMAVGAT----- 285

QY 190 SPGLADNPNIHIAQFSRGATDRGRIKPDVTAPGTFFILSARSSLAPDSSFWMYNSKYAY 249
 Db 286 -----DQNNRASFQYAGL-----DIVAPGVNVQSTYP-----GSTVAS 321

QY 250 MGGTSMATPIVAGNVA-----QLREHFIKNGRITPKPSLIKAALIAGATDVL 297
 Db 322 LNGTSMATPHVAGAAALVKQNPFSWNSVQIRNH-LKN-----TATSLGSLNLT--- 367


```

QY 298 GYPSG 302
Db 368 -YGSQ 371

RESULT 10
PROA_VIBAL STANDARD; PRT; 534 AA.
AC PL6588;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alkaline serine exoprotease A precursor (EC 3.4.21.-).
GN PROA;
OS Vibrio alginolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=663;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89326126; PubMed=2546861;
RT Deane S.M., Robb F.T., Robb S.M., Woods D.R.;
RT "Nucleotide sequence of the Vibrio alginolyticus calcium-dependent,
RT detergent-resistant alkaline serine exoprotease A.";
RL Gene 76:281-288(1989).
CC -|- SIMILARITY: Belongs to peptidase family S8.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; M25499; AAA27550.1;
CC PIR; JSC173; JSC173.
CC HSP; Q93405; IMPT.
CC MEROPS; S08.050; -.
CC InterPro; IPR000209; Peptidase_S8.
CC InterPro; IPR007280; pPC.
CC InterPro; IPR009020; Protease_inhib.
CC Pfam; PF00082; Peptidase_S8; 1.
CC Pfam; PF04151; PPC; 1.
CC PRINTS; PR00723; SUBTILISIN.
CC PROSITE; PS00136; SUBTILASE ASP; 1.
CC PROSITE; PS00137; SUBTILASE HIS; 1.
CC PROSITE; PS00138; SUBTILASE SER; 1.
CC Hydrolase; Serine protease; Zymogen; Signal.
CC SIGNAL 1 21 POTENTIAL.
CC PROPEP 22 141 POTENTIAL.
CC CHAIN 142 534 ALKALINE SERINE EXOPROTEASE A.
CC ACT SITE 180 180 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT SITE 363 363 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC SEQUENCE 534 AA; 55930 MW; 84E96D9C649D4226 CRC64;

Query March 11.7; Score 265; DB 1; Length 534;
Best Local Similarity 26.3; Pred. No. 4.5e-10;
Matches 123; Conservative 65; Mismatches 152; Indels 128; Gaps 24;

QY 7 IVKADVAQ-----NNY--GLYGQGVVAVADTGLDTRDSSMHEAF 46
Db 137 IVSADANQTAIWGLDRIDQENFLDNNYSANFDGTGVTAYVIDTGV-----NNAHVEF 190
QY 47 RGKITYAL-ALGRNNANDPNHGCHTHVAGSVLGNALNKGAPQANLVFQSIM--DSSGGL 103
Db 191 GGRVSGYDFDNDADASDCNGHGHVAGTIGSLY--GVAKNVNLGVRLVSCSGSGT 248
QY 104 GGLSNLNTLPSQAWNAGARHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEG 163
Db 249 SGVIAGVDWVAANA--SGPSVANNSLG---GGQSVALDSAVQSAVQSG-VSFMLAGNSN 302

164 PMSGTISAPGTAKNAITVGATENYRPSFGSLADPNHIAQFSRSGATDRGRIKPDVTAPG 223
303 ADACNYS-PARVATGVTVGST-----TSTDARSFSNWGSC-----VDVFAPG 344
224 TFLSARSSLAPDSFWMYNSKYAYMGSTMATPIVAGNVAQLREHFFKNGITPKPSL 283
345 SQIKSA-----W--YDGGYKTIISGTSMATPHVAG-VAAL--YLOENSSVS--PSQ 387
284 IKAALITAGA-----TDVGLGYPSG-----DQWGRVTLDKSLNVAY 319
388 VEALIVSRASSTKVTTRGSVKNLLYSLTDADCGCGGPDPTDPPEGKLTSGVPVS--- 444
320 VNEATALATQKATYSFQAQAKPLKISLVMTDAPGSTTASYTLVNDLVLVITAPNGQKY 379
445 ---GLSGSGQVAYYYVDVEAGQRLTVQM-----YGGSGDADLYLRF--GAK- 486
380 VGNDFSYPYDNNWDGR-----NNVENFINAPQSGTGTIEVOAYNVPSG 423
487 -----PTLNWDCRPFKYGNNTCTVTSATQSGRYHVMIQSYNSYSG 527

RESULT 11
SUBT_BACS9 STANDARD; PRT; 420 AA.
AC P28842;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Subtilisin precursor (EC 3.4.21.62).
GN SUB1
OS Bacillus sp. (strain TA39).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=29336;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92256481; PubMed=1581352;
RT Narinx E., Davail S., Feller G., Gerdard C.;
RT "Nucleotide and derived amino acid sequence of the subtilisin from
RT the antarctic psychrotroph Bacillus TA39.";
RL Biochim. Biophys. Acta 1131:111-113(1992).
CC -|- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
CC it catalyzes the hydrolysis of proteins and peptide amides.
CC -|- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large uncharged residue
CC in pI. Hydrolyzes peptide amides.
CC -|- COPACITOR: Binds 1 calcium ion per subunit (Potential).
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- MISCELLANEOUS: Still active at temperatures close to 0 degrees
CC Celsius, it has a marked heat lability.
CC -|- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
CC sporulation, and many mutations which block sporulation at early
CC stages affect expression levels of subtilisin. However, subtilisin
CC is not necessary for normal sporulation.
CC -|- SIMILARITY: Belongs to peptidase family S8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; X62369; CAA44227.1; -.
CC PIR; S23407; S23407.
CC HSP; Q99405; IMPT.
CC MEROPS; S08.05A; -.
CC InterPro; IPR000209; Peptidase_S8.
CC Pfam; PF00082; Peptidase_S8; 1.
CC PRINTS; PR00723; SUBTILISIN.
CC PROSITE; PS00136; SUBTILASE ASP; 1.
CC PROSITE; PS00137; SUBTILASE HIS; 1.

```

```
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Sporulation; Serine protease; Zymogen; Metal-binding;
KW Calcium-binding; Signal.
FT SIGNAL 1 26
FT PROPP 27 111
FT CHAIN 112 420
FT ACT_SITE 145 145
FT ACT_SITE 182 182
FT ACT_SITE 360 360
FT ACT_SITE 115 115
FT METAL 154 154
FT METAL 154 154
SQ SEQUENCE 420 AA; 44086 MW; AE4F121BD32B46EC CRC64;

Query Match 11.7%; Score 263.5; DB 1; Length 420;
Best Local Similarity 29.0%; Pred. No. 4.1e-10;
Matches 92; Conservative 46; Mismatches 120; Indels 59; Gaps 15;

QY 21 GQGVAVADTGLDTCGRNDSSMHEAPRGKITAL--YALGRT---NNANDPNHGHTHVAGS 75
DQ 136 GGGINIAVLDGTGNTN-----HPDLNNVQCKDFVTGTTTNNSTCDRQGHGHTVAGS 189
QY 76 VL-----GNALNKGMAPOANLVFQSIM--DSSGGLGGLPSNLTFLFSQAWNAGARIHTN- 127
DQ 190 ALADGGTGVYGVAPADLWAYKVLGDDGGYADDDIAAIRHAGDQATALNTKVWNM 248
QY 128 SWCAPVNGAYTANSRQVDEVRNDMTVLFAAGNECPNSGTISAPGTAKNAITVGATERY 187
DQ 249 SLGSSGESSLITNA---VNYSNKGVLLIAAGNSGPGYSGIPGALVNAVVALEN- 304
QY 188 RPSFGSLADPNPHIAQFSSRG-----ATRDGRIPDVTAPGTFFLSARSLAPDSSF 239
DQ 305 -----KVENCYTRVADFSSRGYSWTDGYAIQKGDV--BISAPGAAYST----- 347
QY 240 WANNYSKYAMNGTSMATFVAGNVAQLREHFNGRITPKPSLKAALIAAGTVDVGLY 299
DQ 348 W--FDGGYATISGTWASPHAGLAAGAKIWAQVPSASNVVDVGELOYRAY---ENDILSGY 402
QY 300 PSG-----DQMGWRVTL 311
DQ 403 YAGYGDGDFASGFGFAIV 419

RESULT 12
SUBB BACLE
ID SUBB BACLE STANDARD; PRT; 269 AA.
AC P29599;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Subtilisin BL (EC 3.4.21.62) (Alkaline protease).
OS Bacillus lentus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1467;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).
RX MEDLINE=93085738; PubMed=1453465;
RA Goddette D.W., Paech C., Yang S.S., Mielenz J.R., Bystroff C.,
RA Wilke M.E., Fletcher R.J.;
RT "The crystal structure of the Bacillus lentus alkaline protease,
RT subtilisin BL, at 1.4-A resolution.";
RL J. Mol. Biol. 228:580-595(1992).
CC -!- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
CC it catalyzes the hydrolysis of proteins and peptide amides.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large uncharged residue
CC in P1. Hydrolyzes peptide amides.
CC -!- COFACTOR: Binds 2 calcium ions per subunit.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
CC sporulation, and many mutations which block sporulation at early
CC stages affect expression levels of subtilisin. However, subtilisin
CC is not necessary for normal sporulation.
CC -!- SIMILARITY: Belongs to peptidase family S8.
```

```
DR MEROPS; S08.003; -.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PRO0723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_SER; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Sporulation; Serine protease; Metal-binding;
KW Calcium-binding; 3D-structure.
FT ACT_SITE 32 32
FT ACT_SITE 62 62
FT ACT_SITE 215 215
FT METAL 2 2
FT METAL 73 73
FT METAL 75 75
FT METAL 77 77
FT METAL 79 79
FT METAL 163 163
FT METAL 165 165
FT METAL 168 168
FT HELIX 6 10
FT TURN 11 12
FT HELIX 13 18
FT TURN 19 20
FT TURN 24 25
FT STRAND 27 32
FT TURN 39 40
FT STRAND 43 48
FT TURN 51 52
FT HELIX 62 71
FT TURN 84 85
FT STRAND 87 92
FT TURN 96 97
FT HELIX 102 114
FT TURN 115 116
FT STRAND 119 122
FT STRAND 126 126
FT HELIX 131 142
FT TURN 143 144
FT STRAND 146 150
FT STRAND 161 161
FT TURN 162 164
FT TURN 166 167
FT STRAND 168 174
FT TURN 176 177
FT STRAND 180 180
FT TURN 182 183
FT TURN 188 189
FT STRAND 192 195
FT STRAND 199 203
FT TURN 204 206
FT STRAND 207 211
FT HELIX 214 231
FT TURN 233 234
FT HELIX 237 246
FT TURN 247 247
FT STRAND 249 249
FT HELIX 254 257
FT TURN 258 259
FT STRAND 261 261
FT HELIX 264 267
FT TURN 268 269
SQ SEQUENCE 269 AA; 26923 MW; B8AFF1A6A9E2676B CRC64;

Query Match 11.5%; Score 259; DB 1; Length 269;
Best Local Similarity 31.5%; Pred. No. 4.5e-10;
Matches 96; Conservative 34; Mismatches 101; Indels 74; Gaps 14;

QY 13 AKNYGLYQQVAVADTGLDTCGRNDSSMHEAPRGKITALYALGRTNANDPNHGHTHV 72
DQ 15 AAHNRGLTSGVKVAVLDTGIST-----HPDLNIRGSGASFVPEP-STODGNHGHTHV 66
```

QY 73 AGSV--LGNALN-KGMAPQANLVFQSIMDSSGGLGGLPSNLNLTFSQANWAGARIHTNSW 129
D6 67 AGTIALANNSIGVLGVAPSAAELVAVKVLGADG--RGAISISIAQGLEWAGNGMGHVALNSL 124
QY 130 GAPVNGAYTANSRQVDEYVYNNDMTVLFAAGNEGPNSTISAPGTAKNAITVGAENYRP 189
D6 125 GSP---SPSATLQAVNSATSRGVLVVAASGNSGASS--ISYPARYANMAVAGAT----- 174
QY 190 SPGLADNPNHIAQFSSRGATROGRKPDVTAGTFTILSARSLAPDSSFWANYNSKIAY 249
D6 175 -----DQNNRASFSQYGAGL-----DIVAFGVNVQSTYP-----GSTYAS 210
QY 250 MGGTSMATPIVAGNVA-----QLREHFINKRGITPKPSLIKAALTAGATDVGL 297
D6 211 LNGTSMATPHVAGAAALVKQNPNSVQIRNH-LKN-----TATSLGSNTL--- 256
QY 298 GYPSG 302
D6 257 -YGSQ 260

RESULT 13
THES_BACSP STANDARD; PRT; 401 AA.
AC Q45670;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Thermophilic serine proteinase precursor (EC 3.4.21.-) (Ak.1
D6 protease).
OS Bacillus sp. (strain AK1)
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95085262; PubMed=7993087;
EA Maciver B., McHale R.H., Saul D.J., Bergquist P.L.;
RT "Cloning and sequencing of a serine proteinase gene from a
thermophilic Bacillus species and its expression in Escherichia
coli.";
RL Appl. Environ. Microbiol. 60:3981-3988 (1994).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.80 ANGSTROMS) OF 122-401.
RX MEDLINE=20057863; PubMed=10589904;
RA Smith C.A., Toogood H.S., Baker H.M., Daniel R.M., Baker E.N.;
RT "Calcium-mediated thermostability in the subtilisin superfamily: the
crystal structure of Bacillus Ak.1 protease at 1.8-A resolution.";
RL J. Mol. Biol. 294:1027-1040 (1999).
CC -|- COPACTOR: Binds 3 calcium ions and 1 sodium ion per subunit.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- MISCELLANEOUS: Has a pH optimum of 8.5, a temperature optimum of
75 degrees Celsius.
CC -|- SIMILARITY: Belongs to peptidase family S8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC
CC EMBL; L29506; AAA63688.1;
DR PIR; I39974; I39974.
DR PDB; 1DBI; 18-NOV-99.
DR MEROPS; S08.009.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR009020; Protease_inhib.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
Signal; 3D-structure.
FT SIGNAL 1 24
FT PROPEP 25 121
FT CHAIN 122 401
FT ACT_SITE 160 160
FT ACT_SITE 193 193
FT ACT_SITE 347 347
FT METAL 126 126
FT METAL 168 168
FT METAL 169 169
FT METAL 171 171
FT METAL 179 179
FT METAL 184 184
FT METAL 186 186
FT METAL 204 204
FT METAL 207 207
FT METAL 209 209
FT METAL 211 211
FT METAL 297 297
FT METAL 300 300
FT METAL 323 323
FT DISULFID 258 260
FT TURN 127 128
FT HELIX 129 132
FT TURN 135 136
FT HELIX 137 139
FT TURN 140 140
FT HELIX 141 144
FT TURN 145 147
FT TURN 152 153
FT STRAND 155 160
FT TURN 165 166
FT TURN 168 173
FT STRAND 174 179
FT TURN 180 183
FT STRAND 184 184
FT HELIX 193 202
FT STRAND 219 224
FT TURN 228 229
FT HELIX 234 246
FT TURN 247 248
FT STRAND 251 254
FT HELIX 263 274
FT TURN 275 276
FT STRAND 278 282
FT STRAND 285 285
FT TURN 298 299
FT STRAND 301 306
FT TURN 308 309
FT TURN 312 312
FT TURN 314 315
FT STRAND 316 316
FT TURN 321 321
FT STRAND 324 327
FT STRAND 331 335
FT TURN 336 338
FT TURN 339 343
FT HELIX 346 362
FT TURN 363 364
FT HELIX 367 376
FT TURN 377 377
FT STRAND 379 379
FT TURN 382 383
FT STRAND 384 384
FT TURN 385 387
FT STRAND 388 388
FT STRAND 391 392
FT HELIX 395 399
FT TURN 401 401
SQ SEQUENCE 401 AA; 42835 MW; 1C736EF4A89F256F CRC64;

```
Query Match 11.3%; Score 255; DB 1; Length 401;
Best Local Similarity 31.7%; Pred. No. 1.4e-09;
Matches 85; Conservative 26; Mismatches 99; Indels 58; Gaps 11;

QY 14 QNNYGLY-----GGQVAVADGLDTRGNDSSNHEAFRGKITALYALGRTNNNDP- 65
DB 137 QNTYTDYANDVTKGSGQBIADTGV-----YTHPLDQKIVKGYDF--VDNDYDPM 188
QY 66 --NGHSTHVAGSVLGNALN----KGMAPQANLVFQSIMDSSGGLGLPSNLTLFSAQWN 119
DB 189 DLNNGHSTHVAGIAAAETNATGIAGWPNTRILAVRALDRNG--SGTSLDIADALIYAD 246
QY 120 AGARIHTNSGAPVNGAYTANSGQVDEYVRNDMTVLFAAGNEGNSGTISAPGTAKNAI 179
DB 247 SGAEVINLSLGC---DCHTTLLENAVNYAANKGSVVAAAGNNG--SSTTEPASVENVI 301
QY 180 TVGATENYRPSFGSLADNPNHIAQFSRGATDRGKIPDVTAPGTFFILSARSLAPDSF 239
DB 302 AVGADVQY-----DRLASTSNYGTW-----VDVAFGVDIIVSTIT----- 336
QY 240 WANYNSKIAYMGTSMTATFIVAGNVAQL 267
DB 337 ----GNRYAYMGSMTASPHVAGLAALL 360

RESULT 14
ID WPA_BACSU STANDARD; PRT; 894 AA.
AC P54423; O06726;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cell wall-associated protease precursor [EC 3.4.21.-] [Contains: Cell
DE wall-associated polypeptides CWBP23 and CWBP52].
GN WPA CR BSU10770.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-54 AND 414-428.
RC STRAIN=168;
RX MEDLINE=97156234; PubMed=9004506;
RA Margot P., Karamata D.;
RT "The wpa gene of Bacillus subtilis 168, expressed during exponential
RT growth, encodes a cell-wall-associated protease.";
RL Microbiology 142:3437-3444(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98015415; PubMed=9353931;
RA Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;
RT "Sequencing of regions downstream of addA (98 degrees) and citG (289
RT degrees) in Bacillus subtilis.";
RL Microbiology 143:3305-3308(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Betero M.G., Besieres P., Bolotin A., Borchert S.,
RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Bruchet S., Bruchet C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dutertre A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrar E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaer-Schanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
```

```
Lee S.M., Levine A., Liu H., Masuda S., Mael C., Medigue C.,
Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
Prasecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Soldo B.,
Sokichi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
Takeuchi M., Takakoshi A., Tanaka T., Takahashi H., Takemaru K.,
Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -1- FUNCTION: NOT YET KNOWN; COULD BE INVOLVED IN PROTEOLYCAN
CC DEGRADATION BY CLEAVAGE OF ITS PEPTIDE BRIDGES.
CC -1- SUBCELLULAR LOCATION: Cell-wall bound.
CC -1- PTM: PROCESSED INTO CWBP23 AND CWBP52.
CC -1- SIMILARITY: Belongs to peptidase family 58.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U58981; AAC25926.1; -
CC EMBL; Y09476; CAA70641.1; -
CC EMBL; Z99109; CAB12917.1; -
CC PIR; F69730; F69730.
CC HSRP; Q45670; IDBI.
CC MEROPS; S08.004; -.
CC Subtilisin; BG11846; wpa.
CC Interpro; IPR000209; Peptidase S8.
CC Pfam; PF00082; Peptidase S8; 1.
CC PRINTS; PR00723; SUBTILISIN.
CC PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
CC PROSITE; PS00137; SUBTILASE_HIS; 1.
CC PROSITE; PS00138; SUBTILASE_SER; 1.
CC Hydrolyase; Serine protease; Cell wall; Zymogen; Signal;
KW Complete proteome.
FT SIGNAL 1 31
FT CHAIN 32 894 CELL WALL-ASSOCIATED PROTEASE.
FT CHAIN 32 ? CWBP23.
FT PROPEP ? 413 POTENTIAL.
FT CHAIN 414 894 CWBP52.
FT ACT_SITE 466 466 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 497 497 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 650 650 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 9 9 V -> A (IN REF. 1).
FT CONFLICT 14 14 L -> I (IN REF. 1).
SQ SEQUENCE 894 AA; 96487 MW; 0F67C353E5F8DBC CRC64;

Query Match 11.1%; Score 250; DB 1; Length 894;
Best Local Similarity 24.3%; Pred. No. 8.1e-09;
Matches 106; Conservative 64; Mismatches 140; Indels 126; Gaps 18;

QY 25 VVAVADTGLDTRGNDSSNHEAFRGKITALYA---LGRTNANDPNHGHTHVAGSVLGNAL 81
DB 457 LIAVVDVGVDSTLAD-----LKGKVRTDIGHNFGVRNNAMDDQGGHTHVAGLIAAQSD 510
QY 82 N----KGMAPQANLVFQSIMDSSGGLGLPSNLTLFSAQWAGARITHNSWGA PNGAY 137
DB 511 NGYSMTGLNKAKEIIPVKVLDSAG--SGDTEQIALGIKYADKGAKEINISLG---GGY 564
QY 138 TANSGQVDEYVRNDMTVLFAAGNEGNSGTISAPGTAKNAITVGATENTRPSFGSLADN 197
DB 565 SRVLEFALKYAADKNVLIATAAASGNDGENA--LSYPASSKYVMVSGAT-----NR 611
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 32.475 Seconds

(without alignment)
4206.909 Million cell updates/sec

Title: US-09-985-689A-5

Perfect score: 2261

Sequence: 1 NDVARGIVKADVAQNNGYLY.....EQVAVNVPSPQRFSLAIVH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvrius:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|--------|-------|--------|--------|---------------------|
| 1 | 2261 | 100.0 | 433 | Q9AQR1 | Q9AQR1 bacillus sp |
| 2 | 2251 | 99.6 | 433 | Q9AQR4 | Q9AQR4 bacillus sp |
| 3 | 2240 | 99.1 | 433 | Q9AQR2 | Q9AQR2 bacillus sp |
| 4 | 2024.5 | 89.5 | 434 | Q9AQR0 | Q9AQR0 bacillus sp |
| 5 | 2002.5 | 88.6 | 639 | Q9AQR3 | Q9AQR3 bacillus sp |
| 6 | 1998.5 | 88.4 | 640 | Q93UV9 | Q93UV9 bacillus sp |
| 7 | 465 | 20.6 | 1825 | Q8T9W1 | Q8T9W1 dictyosteli |
| 8 | 441.5 | 19.5 | 1702 | Q9GYN7 | Q9GYN7 dictyosteli |
| 9 | 422.5 | 18.7 | 654 | Q8U0C9 | Q8U0C9 pyrococcus |
| 10 | 402.5 | 17.8 | 561 | Q8RBJ2 | Q8RBJ2 thermococcus |
| 11 | 395.5 | 17.5 | 1239 | Q9FBZ4 | Q9FBZ4 streptomyces |
| 12 | 377.5 | 16.7 | 430 | Q8ENV1 | Q8ENV1 oceanobacil |
| 13 | 357 | 15.8 | 1253 | Q9FC06 | Q9FC06 streptomyces |
| 14 | 347.5 | 15.4 | 1102 | P95684 | P95684 streptomyces |
| 15 | 345.5 | 15.3 | 412 | Q9AER6 | Q9AER6 thermococcus |
| 16 | 344 | 15.2 | 444 | Q9KBJ7 | Q9KBJ7 bacillus ha |

| | | | | | |
|----|-------|------|------|----|--------|
| 17 | 343.5 | 15.2 | 412 | 16 | Q8RC68 |
| 18 | 343 | 15.2 | 1105 | 2 | Q8KXH6 |
| 19 | 341.5 | 15.1 | 1237 | 2 | Q8GCT4 |
| 20 | 338 | 14.9 | 1220 | 16 | Q9LOA0 |
| 21 | 329 | 14.6 | 1139 | 16 | Q82139 |
| 22 | 328 | 14.5 | 824 | 2 | Q45464 |
| 23 | 328 | 14.5 | 891 | 1 | Q93635 |
| 24 | 328 | 14.5 | 1245 | 16 | Q9RL54 |
| 25 | 327 | 14.5 | 1208 | 16 | Q82B14 |
| 26 | 324 | 14.3 | 435 | 16 | Q8EMJ3 |
| 27 | 315.5 | 14.0 | 442 | 16 | C31788 |
| 28 | 309 | 13.7 | 1398 | 1 | Q9P9L1 |
| 29 | 302.5 | 13.4 | 799 | 16 | Q9KEM1 |
| 30 | 290.5 | 12.8 | 1135 | 1 | Q9P9D1 |
| 31 | 289.5 | 12.8 | 1407 | 16 | Q816G4 |
| 32 | 285 | 12.6 | 1345 | 1 | Q54437 |
| 33 | 284.5 | 12.6 | 1101 | 16 | Q82CF0 |
| 34 | 281.5 | 12.5 | 959 | 16 | Q8PMS7 |
| 35 | 279 | 12.3 | 431 | 2 | Q9S3L6 |
| 36 | 277.5 | 12.3 | 1098 | 16 | Q9L1Z8 |
| 37 | 272.5 | 12.1 | 715 | 2 | P70765 |
| 38 | 271 | 12.0 | 575 | 16 | Q8FNM1 |
| 39 | 269 | 11.9 | 434 | 2 | O54327 |
| 40 | 266.5 | 11.8 | 1571 | 2 | Q8GCM3 |
| 41 | 266 | 11.8 | 530 | 2 | Q8GB52 |
| 42 | 266 | 11.8 | 586 | 16 | Q8PAL8 |
| 43 | 265 | 11.7 | 966 | 16 | Q8PB28 |
| 44 | 264 | 11.7 | 617 | 2 | Q931O4 |
| 45 | 260.5 | 11.5 | 1570 | 16 | Q8E2V6 |

ALIGNMENTS

RESULT 1

| ID | Q9AQR1 | PRELIMINARY; | PRT; | 433 AA. |
|-----|---|--------------|------|---------|
| AC | Q9AQR1; | | | |
| DT | 01-JUN-2001 (TrEMBLrel. 17, Created) | | | |
| DT | 01-JUN-2001 (TrEMBLrel. 17, Last sequence update) | | | |
| DT | 01-JUN-2003 (TrEMBLrel. 24, Last annotation update) | | | |
| DE | Protease (Fragment). | | | |
| GN | PROD. | | | |
| OS | Bacillus sp. S521. | | | |
| OC | Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus. | | | |
| OX | NCBI_TaxID=133780; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=SD521. | | | |
| RX | MEDLINE=20586675; PubMed=1118284; | | | |
| RA | Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H., | | | |
| RA | Horikoshi K.; | | | |
| RT | "Novel oxidatively stable subtilisin-like serine proteases from | | | |
| RT | alkaliphilic Bacillus sp.: enzymatic properties, sequences, and | | | |
| RT | evolutionary relationships." | | | |
| PL | Biochem. Biophys. Res. Commun. 279:313-319(2000). | | | |
| CC | -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9. | | | |
| DR | EMBL; AB046405; BAB21368.1; - | | | |
| DR | HSSP; Q45670; IDBI. | | | |
| DR | GO; GO:0008233; F:peptidase activity; IEA. | | | |
| DR | GO; GO:0004389; F:subtilase activity; IEA. | | | |
| DR | GO; GO:0006508; P:proteolysis and peptidolysis; IEA. | | | |
| DR | InterPro; IPR000205; Peptidase_S8. | | | |
| DR | InterPro; IPR007280; PPC. | | | |
| DR | Pfam; PF00082; Peptidase_S8; 1. | | | |
| DR | Pfam; PF04151; PPC; 1. | | | |
| DR | PRINTS; PR00723; SUBTILISIN. | | | |
| DR | PROSITE; PS00137; SUBTILASE HIS; 1. | | | |
| DR | PROSITE; PS00138; SUBTILASE SER; 1. | | | |
| KW | Hydrolase; Protease; Serine Protease. | | | |
| FT | NON_TER 1 | | | |
| FT | NON_TER 433 | | | |
| FT | NON_TER 433 | | | |
| SEQ | SEQUENCE 433 AA; 45576 MW; 98A2DF18FE660DDC CRC64; | | | |

```
Query Match      100.0%; Score 2261; DB 2; Length 433;
Best Local Similarity 100.0%; Pred. No. 1.8e-124;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
DB 1 NDVARGIVKADVAQNNGYLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
QY 61 NNDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGGLGGLPSNLTLFQAWNA 120
DB 61 NNDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGGLGGLPSNLTLFQAWNA 120
QY 121 GARIHNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
DB 121 GARIHNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSLADPNHIAQFSSRGATRDGRIPKDPVTAPGTFFILSARSSLAPDSSFW 240
DB 181 VGATENYRPSFGSLADPNHIAQFSSRGATRDGRIPKDPVTAPGTFFILSARSSLAPDSSFW 240
QY 241 ANYNSKYAYMGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300
DB 241 ANYNSKYAYMGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300
QY 301 SGQGWGRVTLDSKLVNAVYNEATATGKATYSFQAQAGKPLKISLVMTDAPGSTTAS 360
DB 301 SGQGWGRVTLDSKLVNAVYNEATATGKATYSFQAQAGKPLKISLVMTDAPGSTTAS 360
QY 361 YTLVNDLVLVITAPNGQKYYGNDFSYPYDNNWGRNENVENFINAPQSGTYYTIEVQAYNV 420
DB 361 YTLVNDLVLVITAPNGQKYYGNDFSYPYDNNWGRNENVENFINAPQSGTYYTIEVQAYNV 420
QY 421 PSGQRFSLAIVH 433
DB 421 PSGQRFSLAIVH 433

RESULT 2
Q9AQR4
ID Q9AQR4 PRELIMINARY; PRT; 433 AA.
AC Q9AQR4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protease (Fragment).
GN PROA.
OS Bacillus sp. D6.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=127889;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D6;
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RL evolutionary relationships."
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR EMBL; AB046402; BAB21265.1; -.
DR HSSP; Q45670; 1DBI.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HTS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.

KW Hydrolase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 433
SQ SEQUENCE 433 AA; 45636 MW; 52087EOA2516107F CRC64;

Query Match      99.6%; Score 2251; DB 2; Length 433;
Best Local Similarity 99.3%; Pred. No. 7.1e-124;
Matches 430; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
DB 1 NDVARGIVKADVAQNNGYLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
QY 61 NNDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGGLGGLPSNLTLFQAWNA 120
DB 61 NNDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGGLGGLPSNLTLFQAWNA 120
QY 121 GARIHNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
DB 121 GARIHNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSLADPNHIAQFSSRGATRDGRIPKDPVTAPGTFFILSARSSLAPDSSFW 240
DB 181 VGATENYRPSFGSLADPNHIAQFSSRGATRDGRIPKDPVTAPGTFFILSARSSLAPDSSFW 240
QY 241 ANYNSKYAYMGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300
DB 241 ANYNSKYAYMGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300
QY 301 SGQGWGRVTLDSKLVNAVYNEATATGKATYSFQAQAGKPLKISLVMTDAPGSTTAS 360
DB 301 SGQGWGRVTLDSKLVNAVYNEATATGKATYSFQAQAGKPLKISLVMTDAPGSTTAS 360
QY 361 YTLVNDLVLVITAPNGQKYYGNDFSYPYDNNWGRNENVENFINAPQSGTYYTIEVQAYNV 420
DB 361 YTLVNDLVLVITAPNGQKYYGNDFSYPYDNNWGRNENVENFINAPQSGTYYTIEVQAYNV 420
QY 421 PSGQRFSLAIVH 433
DB 421 PSGQRFSLAIVH 433

RESULT 3
Q9AQR2
ID Q9AQR2 PRELIMINARY; PRT; 433 AA.
AC Q9AQR2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protease (Fragment).
GN PROC.
OS Bacillus sp. Y.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=133779;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y;
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RL evolutionary relationships."
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR EMBL; AB046404; BAB21267.1; -.
DR HSSP; Q45670; 1DBI.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptidase_S8; 1.
```



```
DR Pfam: PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Protease; Serine protease.
FT NON_TER 433 433
SQ SEQUENCE 433 AA; 45587 MW; B81291A803C775AE CRC64;

Query Match
Best Local Similarity 99.1%; Score 2240; DB 2; Length 433;
Matches 428; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYLGQGVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 60
Db 1 NDVARGIVKADVAQNNGYLGQGVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 60
QY 61 NNDPNGHGTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGPLSNLTLFQAWNA 120
Db 61 NASDPNGHGTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGPLSNLTLFQAWNA 120
QY 121 GARIHTNSGAPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAIT 180
Db 121 GARIHTNSGAPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGLANPNHIAQFSRGATRDGRIPKPDVTPAGTFTILSARSLAPDSSF 240
Db 181 VGATENYRPSFGLANPNHIAQFSRGATRDGRIPKPDVTPAGTFTILSARSLAPDSSF 240
QY 241 ANYNSKYAYVGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGY 300
Db 241 ANYNSKYAYVGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGY 300
QY 301 SDQGGWGRVTLNLSLVAVYNEATATLQKATYSFQAQAGKPLKISLWTDAPGSTTA 360
Db 301 NGDQGGWGRVTLNLSLVAVYNEATATLQKATYSFQAQAGKPLKISLWTDAPGSTTA 360
QY 361 YTLVNDLDELVTAPNGQKYVGNDFSYPDNNNDGRNNVENFINAPQSGTYTIEVQAYN 420
Db 361 YTLVNDLDELVTAPNGQKYVGNDFSYPDNNNDGRNNVENFINAPQSGTYTIEVQAYN 420
QY 421 PSQGFQFSLAIVH 433
Db 421 PSQGFQFSLAIVH 433

RESULT 4
Q9AQR0 PRELIMINARY; PRT; 434 AA.
AC Q9AQR0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protease (Fragment).
GN PROE.
OS Bacillus sp. NV1.
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
OX NCBI_TaxID=133781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NV1;
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.; enzymatic properties, sequences, and
RT evolutionary relationships."
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR ENBL; AB046406; BAB21269.1; -
DR HSSP; P00782; 1SUP.
DR DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
```

```
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Protease; Serine protease.
FT NON_TER 434 434
SQ SEQUENCE 434 AA; 45294 MW; 83517EDDB74125D2 CRC64;

Query Match
Best Local Similarity 89.5%; Score 2024.5; DB 2; Length 434;
Matches 384; Conservative 28; Mismatches 21; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNGYLGQGVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 60
Db 1 NDVARGIVKADVAQNNGYLGQGVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 60
QY 61 NNDPNGHGTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGPLSNLTLFQAWNA 119
Db 61 NNDPNGHGTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGPLSNLTLFQAWNA 120
QY 120 AGARIHTNSGAPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAI 179
Db 121 AGARIHTNSGAPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAI 180
QY 180 TVGATENYRPSFGLANPNHIAQFSRGATRDGRIPKPDVTPAGTFTILSARSLAPDSSF 239
Db 181 TVGATENYRPSFGLANPNHIAQFSRGATRDGRIPKPDVTPAGTFTILSARSLAPDSSF 240
QY 240 WANYNSKYAYVGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGY 299
Db 241 WANDHSDKYAYVGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGY 300
QY 300 PSQGGWGRVTLNLSLVAVYNEATATLQKATYSFQAQAGKPLKISLWTDAPGSTTA 359
Db 301 PSQGGWGRVTLNLSLVAVYNEATATLQKATYSFQAQAGKPLKISLWTDAPGSTTA 360
QY 360 SYTLVNDLDELVTAPNGQKYVGNDFSYPDNNNDGRNNVENFINAPQSGTYTIEVQAYN 419
Db 361 SVTLVNDLDELVTAPNGQKYVGNDFSYPDNNNDGRNNVENFINAPQSGTYTIEVQAYN 420
QY 420 VPSCGFQFSLAIVH 433
Db 421 VPQGFQFSLAIVN 434

RESULT 5
Q9AQR3 PRELIMINARY; PRT; 639 AA.
AC Q9AQR3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protease.
GN PROA.
OS Bacillus sp. 9860.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=133778;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9860;
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.; enzymatic properties, sequences, and
RT evolutionary relationships."
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR ENBL; AB046403; BAB21266.2; -
DR DR
```


RESULT 8

```
RESULT 9
Q8UBJ2 PRELIMINARY; PRT; 561 AA.
ID AC Q8UBJ2
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Subtilisin-like serine proteases.
GN APRE2 OR TTE0824
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
CX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; AE013049; AAM24081.1;
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR002029; Peptidase_S8.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Protease; Complete proteome.
SQ SEQUENCE 561 AA; 59968 MW; BA9C52F7083A18 CRC64;

Query Match 17.8%; Score 402.5; DB 16; Length 561;
Best Local Similarity 31.0%; Pred. No. 1.3e-15;
Matches 142; Conservative 63; Mismatches 160; Indels 93; Gaps 19;

QY 16 NYGLYGGQVVAVADTGLDTRNDSSMHEAFRGKITALVALGRNTNAN--DPNGHG 69
DB 152 NLGYDGGSGITIGITDGD-----ASHPDLOGKV-----IGWDFVNGRSYPYDDHGHG 200
QY 70 THVAGSVLG-----NALNKGMAPOANLVFQSIW--DSSGSLGLPSNLNLTFSQAWNAGA 122
DB 201 THVASIAAGTGAASNGKYKVAQKLAGIKVLGADGSGSISTIIKGVWAVDNKDKYGI 260
QY 123 RIHTNSKGA PVNGAYTANSRQVDEYVRNNDMTLVFAAGNEGPNSTISAPGTAKNAITVG 182
DB 261 KVINLSLGSQSSDGTDLASQAVNAAMDAGLVVVAAGNSGPNKYITIGSPAAASKVITVG 320
QY 183 ATENYRPSFGLADNPNHIAQFSRGATRDGRIPKDVTPAGTFLSARGLAPDSFWAN 242
DB 321 AVDKY-----DVIITFSRGPTADGRKPEVAPGNWIIAARAS---GTSMGQP 366
QY 243 YNSKYAYMGTSMTPTVAGNVAQLREHFIKRGITP---KPSLIKAALIAGATDVG-LG 298
DB 367 INDYTTAAGTSMATPHVAGIAALLQ---AHPSTWPKVKLTALITADIKVPDEIADIA 423
QY 299 YPSGDQGWGVTLDKSLNVAVNEATALATGQKA-----TYSFQAQAGKPLKISLVWTD 353
DB 424 Y-----GAGRNAYKAN--YDNFVKLVFTGYVANKGSQTHQFVIGASFTVATLWDNA 476
QY 354 PGSTTASYTLVDLITAPNGKQYVGNDFSPYDNNWDRNENNVFNAPQSGTYTI 413
DB 477 N-----SDLDLVLPNGQV---DYSY-----TAYYGFKEKVGYNPTDGTWTI 517
QY 414 EVQAYNVPSGPFQSFSLAIV 432
DB 518 KWSY---SGSANYQVDVW 533

RESULT 10
Q8RBJ2
```

516 TGTQEQETITLPSQTGTGYVVKVYSR-GSGNYFFDL5 552

Db

RESULT 11

Q9FBZ4 PRELIMINARY; PRT; 1239 AA.

AC Q9FBZ4

DT 01-MAR-2001 (TremBLrel. 16, Created)

DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)

DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)

DE Putative secreted peptidase.

GN SC07188 OR SC8A11.16C

OS Streptomyces coelicolor.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomyces.

OC NCBI_TaxID=1902;

RN [1]

RN SEQUENCE FROM N.A.

RP STRAIN=A3(2);

RC Saunders D.C., Harris D.,

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RN SEQUENCE FROM N.A.

RP STRAIN=A3(2);

RC MEDLINE=97000951; PubMed=8843436;

RX Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,

RA Kinashi H., Hopwood D.A.;

RT "A set of ordered cosmids and a detailed genetic and physical map for

RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL Mol. Microbiol. 21:77-96(1996).

RN [4]

RN SEQUENCE FROM N.A.

RP STRAIN=A3(2);

RC MEDLINE=97000951; PubMed=8843436;

RX Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,

RA Kinashi H., Hopwood D.A.;

RT "A set of ordered cosmids and a detailed genetic and physical map for

RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL Mol. Microbiol. 21:77-96(1996).

RN [4]

RN SEQUENCE FROM N.A.

RP STRAIN=A3(2) / M145;

RC MEDLINE=21996410; PubMed=12000953;

RX Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,

RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,

RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,

RA Cronin A., Fraser A., Goble A., Hidaigo J., Hornsby T., Howarth S.,

RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,

RA Rabinowitz B., Rajandream M.A., Rutherford K., Rutter S.,

RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,

RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,

RA Hopwood D.A.;

RT "Complete genome sequence of the model actinomycete Streptomyces

RT coelicolor A3(2).";

RL Nature 417:141-147(2002).

RL EMBL; ALJ93130; CAC01588.1; -.

DR HSSP; Q99405; IMPT.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0004289; F:subtilase activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR003137; FA.

DR InterPro; IPR000209; Peptidase_S8.

DR Pfam; PF02225; PA; 1.

DR Pfam; PF00082; Peptidase_S8; 1.

DR PRINTS; PR00723; SUBTILISIN.

DR PROSITE; PS00136; SUBTILASE_ASP; 1.

DR PROSITE; PS00137; SUBTILASE_HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.

DR PROSITE; PS00139; SUBTILASE_SER; 1.

KW Complete proteome.

SW SEQUENCE 1239 AA; 128505 MW; 8F5E9AC6EB1260A CRC64;

Query Match 17.5%; Score 395.5; DB 16; Length 1239;

Best Local Similarity 30.6%; Pred. No. 9.6e-15;

Matches 141; Conservative 55; Mismatches 172; Indels 93; Gaps 16;

QY 8 VKADVAQNNY-----GLYGGQVAVADTGLDTRNDSSMHEAFRGKITALYALG 57

Db

219 VEADLADSTAQIGAPRAWAGNTQGGVEVAVLDTGVDAG-----HPLADRIARQSFV 272

QY 58 RTNNANDPENGHTHVAGSVLGNAL-----NKGMAPQANLVFQSTMDSSGGGLGGLPSNLNT 112

Db 273 PDENTDRDGRGHTVASTIAGTGAASAGKEKGAVAPGARLSTGKVLDS-GRGQISWTLAA 331

QY 113 LFSQAWNAGARIHTNSWGA-PVNGAYTANSQVDEYVNNDMTVLFAAGNEGPNSTGISA 171

Db 332 MEWAVERHAKIVNNSLGSQSDGSPMSRAVDRLSAQTGALFVVAAGN-GGEAGSIGA 390

QY 172 PGTAKNNAITVGATENYRPSFGSLADNPNHIAQFSRSGATRDGRIPKDPVTAPGTFFILSARS 231

Db 391 PGVATSAITVGA-----VDATDTLAPFSSQGRVVDGALKPBITAPGVGILAA-- 437

QY 232 SLAPDSSFWANYSKYVMGTSMATPIVAGNVLAQLREHTKNGRIFPKSLIKAAL--- 288

Db 438 -----NSSFAAGNGAYQLSGTSMATPHVGAALL-----AAARPDLSSGALKDV 484

QY 289 IAGATDVGLGYPGQDGMGRVTLDKSLNVAVYNNATALATQK-----ATYGFQQAQ 341

Db 485 LASSSHRTFRYDAFQAGSRVDVDAVRAGVYASATAPGSSPGPVRLVYTTNTGAA 544

QY 342 KPLKISLVWTDAPGST---TASYTLV---NDLDLVIT-----ANGQKYVGNDSFYSDN 390

Db 545 VTLELSVAATHAPEGVFRLSARVTPAHGTADVTLTIDGSGSAGRAYSQILA---T 600

QY 391 NWDGRNNVENFINAPQSGTITIEVQAYNVPSGPQFSLAI 431

Db 601 DADARN-----VAHTAVSAGPVRRKLTIV 623

RESULT 12

Q8ENV1 PRELIMINARY; PRT; 430 AA.

AC Q8ENV1

DT 01-MAR-2003 (TremBLrel. 23, Created)

DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)

DE Intracellular alkaline serine proteinase.

GN OB2375.

OS Oceanobacillus iheyensis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.

OX NCBI_TaxID=182710;

RN [1]

RN SEQUENCE FROM N.A.

RP STRAIN=HTE831 / DSM 14371 / JCM 11309;

RC MEDLINE=2220767; PubMed=12235376;

RA Takami H., Takaki Y., Uchiyama I.;

RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya

RT Ridge and its unexpected adaptive capabilities to extreme

RT environments.";

RL Nucleic Acids Res. 30:3927-3935(2002).

DR EMBL; AP004601; BAC14331.1; -.

DR GO; GO:0004289; F:subtilase activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR000209; Peptidase_S8.

DR Pfam; PF00082; Peptidase_S8; 1.

DR PRINTS; PR00723; SUBTILISIN.

DR PROSITE; PS00136; SUBTILASE_ASP; 1.

DR PROSITE; PS00137; SUBTILASE_HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.

KW Complete proteome.

SW SEQUENCE 430 AA; 45838 MW; 6D09A99BBC1E310F CRC64;

Query Match 16.7%; Score 377.5; DB 16; Length 430;

Best Local Similarity 34.5%; Pred. No. 2.7e-14;

Matches 119; Conservative 48; Mismatches 119; Indels 59; Gaps 16;

QY 2 DVARGTGVADVAQNNYGLYGGQVAVADTGLDTRNDSSMHEAFRGKIT--ALYALGRT 59

Db 121 DTASSINADVLKES-GTGGQSTTAVDTGTHP-----HEDLEGIIFADFVGQT 172

QY 60 NNANDPENGHTHVAGSVLGN-ALN----KGMAPQANLVFQSTMDSSGGGLGGLPS----- 108

[illegible]

```

Db 173 EPYDD-NHGTGHCAGDAAGNALSOGYQCPAPDANLVGVKLNKGTGS-GSLSTVIEGID 230
QY 109 -----NINLTFSQAWNAGARIHTNSGAPVNGAYTANSQVDEYVRNNDMTVLPAA 159
Db 231 WCIONOSKNINIL-----SUSLGSDATEPAGGPPVNAV-----ETAWDNGMVVCAA 279
QY 160 GNEGPNSTISAFGTAKNAITVGTATNYPFSGSLADNPNHIAQFSSRGATDGRKPDV 219
Db 280 GNSGPGDKTVGSPGISPKVITVGAADDNNTAERS-----DDSAFSSRGPTIDGLTKENL 335
QY 220 TAPGTILSARS--SLAPSSFWANNYSKYAYMGTSMTATPIVAGNVAQLREHFTKNRGI 277
Db 336 LTGPGDIVSLRAFGSFDIKNTKSARVGSNYISLSGTSMTATPICAGIVAQLLQ---SDSSL 392
QY 278 TPKEPSLKAALIAGATDVLGVPESGQGGGRVTLDKSLNVAYVNE 322
Db 393 T--PNQVKEKLMEACQD--LGQSPNVQAGYL--NAANLINE 430

RESULT 13
Q9FC06 PRELIMINARY; PRT; 1253 AA.
ID Q9FC06
AC Q9FC06;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative secreted peptidase.
GN SCO7176 OR SC8A11.04C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Denapaita D., Eichner A., Cullum J.,
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb streptomycetes coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
RL EMBL; AL939130; CAC01576.1; -.
DR HSSP; Q99405; IMPI.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

```

```

DR InterPro; IPR003137; PA.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF02235; PA; 1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Complete proteome.
SQ SEQUENCE 1253 AA; 130971 MW; AA69B417BFEEDB89 CRC64;

Query Match 15.8%; Score 357; DB 16; Length 1253;
Best Local Similarity 30.0%; Pred. No. 1.7e-12;
Matches 143; Conservative 53; Mismatches 165; Indels 116; Gaps 23;

QY 18 GLYQGGQWVAVDTGLDTRNDSSMEHAFRGKITALYALGRTNNANDPNHGHTHVAGSVL 77
Db 235 GNTGEGVAVLDTGVDA-----HPDFAGRTAATASFVDPQDVTDRNGHGHVASTVA 288
QY 78 G-----NALNKGMAPOANLVFQSIMDSSGGLGPNLNTLPSQAM---NAGARIHTNSW 129
Db 289 GTGAASGGVEKGVAPGASLHIGKVLDSG--SQDSNV--LAGMEWAVRDQHAKIVSMSL 344
QY 130 G-APVNGAYTANSRQVDEYVRNNDMTVLPAAGNEGPNSTISAPGTAKNAITVGATENYR 188
Db 345 GDSFTDGT-DPLSEAVNNLSAETGALFVVAAGNSGPEATVGTAAADAALITVAVNGPG 403
QY 189 PSFGSLADNPNHIAQFSSRG-ATRDGRIKPDVTAPGTFTILSARSLAPSSFWANNYSKY 247
Db 404 KGVDQLAD-----FSSRGRVGDNAVKPDLTAPGVGVLAARSRYAPEG-----EGAY 450
QY 248 AYMGTSMTATPIVAGNVAQL-RSHFIKNGRITPKP-----SLIKAAALIAGATDVLGVP 302
Db 451 QSLSGTSMATPHVAGAAALLAAEH-----PDWTGORLEALV-GTTAGTQRFSPF 499
QY 303 DQGWGRVTLDKSLNVAYVNEATALTG-----OKATYSFQAGKPLK 345
Db 500 DAGSGRV-----DVAARVSTLLASGDAFAQAHPYTPGTQVRRDVTVTSNGPA--PVA 551
QY 346 ISLVWTD-----PGSTTASYTLVNDLD-----LVITAPNGQK 378
Db 552 LDIALSPAEPLPEGLTLESAQVTVPAHGTSVGVITHLDAEDNGAYATRLVASGADGAV 611
QY 379 YVGNDFSYPDNNWGNVNVFINA-----POSTTYTIEVOAYNVSPGQRFSL 429
Db 612 LART----PVGVNKEGR--ATLALAKDHDHDKPLSGTVILKQVERN--TAPKVYSV 660

RESULT 14
P95684
ID P95684 PRELIMINARY; PRT; 1102 AA.
AC P95684;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Subtilisin-like protease.
OS Streptomyces albobacillus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S-3253;
RX MEDLINE=97144528; PubMed=8990295;
RA Suzuki M., Taguchi S., Yamada S., Kojima S., Miura K., Momose H.;
RT "A novel member of the subtilisin-like protease family from
RT Streptomyces albobacillus.";
RL J. Bacteriol. 179:430-438(1997).
CC !- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR EMBL; D83672; BAA12040.1; -.
DR HSSP; P00782; 2SBI.
DR MEROPS; S08.069; -.
DR GO; GO:0008233; F:peptidase activity; IEA.

```

```
DR GO: 0004289; F: subtilase activity; IEA.
DR GO: 0006508; P: proteolysis and peptidolysis; IEA.
DR InterPro: IPR002860; GH_BNR.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF02012; BNR; 2.
DR PRINTS: PR0082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DR PROSITE: PS00139; SUBTILASE_SER; 1.
DR Hydrolase; Protease; Serine protease.
KW SEQUENCE 1102 AA; 1412 MW; F9E4AD2590FE559E CRC64;

Query Match      15.4%; Score 347.5; DB 2; Length 1102;
Best Local Similarity 29.7%; Pred. No. 5.3e-12;
Matches 130; Conservative 52; Mismatches 181; Indels 75; Gaps 16;

QY 18 GLYGQGVVAVADTGLDGTGRNDSSMHEAFPRKITYALGRNTNANDPNHGHTHVAGSVL 77
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
212 GYDGRGVXIAVLDTGVD-----ATHPDLKQVTSKNTSAPTITGVDVVGHTHVASIAA 265
QY 78 GNALN-----KGMAPQANLVFQSIIMDSGGGLGLPSNLNTLFSQAWNA--GARITHNSWG 130
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
266 GTCAQSKTKYKGVAVGAKILNGVLDDA----GFGDDSGILAGMEWAAQAQADIVNWSLG 321
QY 131 A-----PVNGAYTANRQVDEYVRNNDMTLVLPAGNEGPNSTISAPGTAKNAITVGA 183
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
322 GMDTPTDPLEAA-----VDKLSAEKGIILFAIAAGNEGPOS--IGSPGADSALTVGA 372
QY 184 TENYPSFGSLADNPNHIAQESSRG-ATRDGRIKPDVTAPGTFILSARSSILAPSSFWAN 242
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
373 -----VDDKDKLADFSSTGPRLGDAVPELTPAGVDITAAASAKNDIAKBUVE 421
QY 243 YNSKYAYMGTSMATPIVAGNVAQUREHFIKNRGITP--KPSLIKAALIAGATDVGLGYP 300
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
422 KPAGYMTISGTSMAPEHVAGAAALIKQOH-----PEWKYAEKLGALTASTKDG--KYT 472
QY 301 SDQCGWVTLDKSLNVAIVNEATATAG-----OKATYSFQAQAKPLKIS 347
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
473 PEEQSGRQVQMDKATQTVIAEPVSLSFQVQWPHADDPKVTYKLTLYRNLTGTDVTLKLT 532
QY 348 LWMTDAPG-STTASYTLVNDLDTVITAPNGQKYGVNDFSYFYDNNWGRNNVENVFINAP 406
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
533 STATGPKGAAPAGFTLGASTLTVPA-NGTASVDVTDATRLGGAVDGTYSAYVVATGAG 591
QY 407 QS-----GYTTEVQAYNV 420
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
592 QSVRTAAAVEREVSINV 609

RESULT 15
Q9AER6 PRELIMINARY; PRT; 412 AA.
AC Q9AER6
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Thermicin (subtilisin-like protease thermicin).
OS Thermoanaerobacter yonsei.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=111519;
RN [1]
RP SEQUENCE FROM N.A.
RA Jang H.J., Kim B.C., Pyun Y.R., Kim Y.S.;
RT "A novel subtilisin-like serine protease from Thermoanaerobacter
RT yonseiensis K2-1: cloning, expression and biochemical properties.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Jang H.J., Kim B.C., Pyun Y.R., Kim Y.S.;
RT "Subtilisin-like protease, thermicin, from Thermoanaerobacter
RT yonsei.";
```

Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
EMBL; AY028704; AA027733.1; -;
EMBL; AF305633; AA09366.1; -;
HSP; Q45670; 1DBI.

GO: 0008233; F: peptidase activity; IEA.
GO: 0004289; F: subtilase activity; IEA.
GO: 0006508; P: proteolysis and peptidolysis; IEA.
InterPro: IPR000209; Peptidase_S8.
Pfam: PF00082; Peptidase_S8; 1.
PRINTS: PR00723; SUBTILISIN.
PROSITE: PS00136; SUBTILASE ASP; 1.
PROSITE: PS00137; SUBTILASE_HIS; 1.
PROSITE: PS00138; SUBTILASE_SER; 1.
Hydrolase; Protease; Serine protease.
KW SEQUENCE 412 AA; 44503 MW; E3C6A0P81B1A1D47 CRC64;

Query Match 15.3%; Score 345.5; DB 2; Length 412;
Best Local Similarity 35.5%; Pred. No. 1.9e-12;
Matches 116; Conservative 39; Mismatches 113; Indels 59; Gaps 17;
QY 15 NNLYGQGVVAVADTGLDGTGRNDSSMHEAF---RGKITALY-ALGRTNNDPNHGHT 70
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 NDLYGTEKGIITAFDGTG-----YHPDFTKPNRIIAFYDVVNGKKQPYDDNGHGT 168
QY 71 HVAGSVLGN--ALN---KGMAPQANLVFQSIIMDSGGGLGLPSNLNTLFSQAW-----N 119
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
169 HVAGDAAGNGYASNGKYKGVAPKANIVAVKVLDSY----GRGSSSDILAGMQWLDNKEK 224
QY 120 AGARIHTNSWG-APVNGAYTANRQVDEYVR-----NNDMTVLPAAGNEGPNSTISAP 172
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
225 YNIRIVSLSIGETPALPTF-----LDPLVRGVDTLWKNGIIVVVAAGNSGNGYNSITSP 278
QY 173 GTAKNAITVGATENYR-PSFGSLADNPNHIAQESSRGATRDGRIKPDVTAPGTFILSARS 231
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
279 GTSRNAITVGAVDDKETPD1-----EDDEVAKFSGRGGPY--LYKPDVVVAPGVKIVSTAS 331
QY 232 SLAPDSSFWANTNSKYAYMGTSMATPIVAGNVAQUREHFIKNRGITPKPSLIKAAALAG 291
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 GNVFFGADEIMINKPYSATGTSMATPMVAGAVALLLE---KNSRLTNVE--IKNLIKTT 386
QY 292 AT---DVGLGYPGSDGQWGRVTLDKSL 315
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
387 ATKINEAGL---WTQSGMINIEAL 409

Search completed: March 31, 2004, 16:09:00
Job time : 34.475 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 47.2392 Seconds
(without alignments)
2595.843 Million cell updates/sec

Title: US-09-985-689A-6

Perfect score: 2250
Sequence: 1 NDVARGIVKADVAQSSYGLY.....EVQAYNVPGQAPSLAIVN 434

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_28Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|--------|-------|--------|-------|-------------|
| 1 | 2250 | 100.0 | 434 | 5 | AAW50085 |
| 2 | 2143 | 95.2 | 434 | 5 | AAW50080 |
| 3 | 2143 | 95.2 | 640 | 2 | AAV17090 |
| 4 | 2141 | 95.2 | 640 | 2 | AAV17091 |
| 5 | 2130 | 94.7 | 434 | 5 | AAW50081 |
| 6 | 2122 | 94.3 | 639 | 2 | AAV17089 |
| 7 | 2116.5 | 94.1 | 433 | 5 | AAW50086 |
| 8 | 2116.5 | 94.1 | 641 | 2 | AAW89547 |
| 9 | 2082 | 92.5 | 639 | 2 | AAV17087 |
| 10 | 2082 | 92.5 | 640 | 2 | AAV17088 |
| 11 | 2024.5 | 90.0 | 433 | 5 | AAW50084 |
| 12 | 2020.5 | 89.8 | 433 | 5 | AAW50082 |
| 13 | 2016.5 | 89.6 | 636 | 2 | AAW89548 |
| 14 | 2010.5 | 89.4 | 433 | 2 | AAW61495 |
| 15 | 2010.5 | 89.4 | 433 | 2 | AAW95698 |
| 16 | 2010.5 | 89.4 | 433 | 3 | AAV69207 |
| 17 | 2010.5 | 89.4 | 433 | 3 | AAV69207 |
| 18 | 2010.5 | 89.4 | 433 | 3 | AAV44619 |
| 19 | 2005.5 | 89.1 | 433 | 5 | AAW50083 |
| 20 | 1973 | 87.7 | 434 | 5 | AAW50090 |
| 21 | 1562.5 | 69.4 | 345 | 2 | AAW62230 |
| 22 | 1562.5 | 69.4 | 345 | 2 | AAW21654 |
| 23 | 450.5 | 20.0 | 659 | 2 | AAW24121 |
| 24 | 450.5 | 20.0 | 659 | 2 | AAW94840 |
| 25 | 404 | 18.0 | 412 | 2 | AAW94836 |

| | | | | | |
|----|-------|------|-------|---|----------|
| 26 | 404 | 18.0 | 522 | 2 | AAW24122 |
| 27 | 404 | 18.0 | 522 | 2 | AAW94838 |
| 28 | 404 | 18.0 | 654 | 2 | AAW24129 |
| 29 | 404 | 18.0 | 654 | 2 | AAW94841 |
| 30 | 403 | 17.9 | 659 | 2 | AAW24123 |
| 31 | 376 | 16.7 | 545 | 4 | ABW9483 |
| 32 | 357 | 15.9 | 1079 | 6 | ABW81180 |
| 33 | 357 | 15.9 | 1079 | 6 | ABW07391 |
| 34 | 355 | 15.8 | 520 | 2 | AAW13666 |
| 35 | 355 | 15.8 | 734 | 2 | AAW13667 |
| 36 | 355 | 15.8 | 823 | 2 | AAW13668 |
| 37 | 339 | 15.1 | 1237 | 6 | ABU11343 |
| 38 | 318.5 | 14.2 | 806 | 2 | AAW27481 |
| 39 | 316.5 | 14.1 | 903 | 2 | AAW87007 |
| 40 | 316.5 | 14.1 | 1398 | 2 | AAW87008 |
| 41 | 316.5 | 14.1 | 1398 | 2 | AAW24124 |
| 42 | 316.5 | 14.1 | 1398 | 2 | AAW94839 |
| 43 | 302.5 | 13.4 | 699 | 2 | AAV08471 |
| 44 | 297.5 | 13.2 | 519 | 6 | ABP76735 |
| 45 | 297.5 | 13.2 | 19938 | 6 | ABP76678 |

ALIGNMENTS

RESULT 1

AAW50085
ID AAW50085 standard; protein; 434 AA.

XX AC AAW50085;

XX AC (first entry)

DT 12-AUG-2002 (first entry)

DE Bacillus sp alkaline protease protein A-1 fragment.

XX Alkaline protease; detergent; laundry; bleaching; dishwasher.

XX Bacillus sp.

XX EPI2092333-A2.

XX PD 29-MAY-2002.

XX PF 22-NOV-2001; 2001EP-00127851.

XX PR 22-NOV-2000; 2000JP-00355166.

XX PR 12-APR-2001; 2001JP-00114048.

XX (KAOS) KAO CORP.

XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;

XX Okuda M, Saeki K;

XX WPI; 2002-437518/47.

XX New modified alkaline proteases useful in detergent compositions.

XX Claim 5; Page 18-19; 25pp; English.

XX This invention describes novel Bacillus sp. alkaline proteases useful in detergent compositions, especially in laundry, bleaching or automatic dishwasher detergents. The novel proteases have an increased agency % (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This sequence represents a fragment of the alkaline protease A-1 from Bacillus sp NCIB12289 described in the method of the invention

XX SQ Sequence 434 AA;

Query Match 100.0%; Score 2250; DB'S; Length 434;

Best Local Similarity 100.0%; Pred. No. 1.3e-160;

Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 NDVARGIVKADVAQSSYGLYGGQVAVADTGLDGRNDSMHEAFRGKTIAYALGRTN 60

Db 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDGRNDSSMHEAFRGKITAIYALGRTN 60
 QY 61 NNDPENGHGHVAGSVLNGTGNKGMAPQANLVFQSVMDNSGGLGLPSNVSTLFSQAYS 120
 Db 61 NNDPENGHGHVAGSVLNGTGNKGMAPQANLVFQSVMDNSGGLGLPSNVSTLFSQAYS 120
 QY 121 AGARIHTNSWGAPVNGAYTTDSRNVDYVVRKNDMAVLFAAGNEGPNNGGTISAPGTAKNAI 180
 Db 121 AGARIHTNSWGAPVNGAYTTDSRNVDYVVRKNDMAVLFAAGNEGPNNGGTISAPGTAKNAI 180
 QY 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTIFLSARSSLAPDSSF 240
 Db 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTIFLSARSSLAPDSSF 240
 QY 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKRGITPKPSLLKAALIAGATDIGLY 300
 Db 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKRGITPKPSLLKAALIAGATDIGLY 300
 QY 301 PSGNQGWGRVTLDKSLNVAFVNETSSLTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 360
 Db 301 PSGNQGWGRVTLDKSLNVAFVNETSSLTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 360
 QY 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGDGRNNVNFVINAPOSQGTYYVEVOAYN 420
 Db 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGDGRNNVNFVINAPOSQGTYYVEVOAYN 420
 QY 421 VPOGPQAFSLAIYN 434
 Db 421 VPOGPQAFSLAIYN 434

RESULT 2

AAM50080
 ID AAM50080 standard; protein; 434 AA.

AC AAM50080;

XX 12-AUG-2002 (first entry)
 DT

DE Bacillus sp KSM-KP43 alkaline protease protein fragment.

XX Alkaline protease; detergent; laundry; bleaching; dishwasher.

XX Bacillus sp.

XX EP1209233-A2.

XX 29-MAY-2002.

XX 22-NOV-2001; 2001EP-00127851.
 XX

XX 22-NOV-2000; 2000JP-00355166.

XX 12-APR-2001; 2001JP-00114048.

XX (KAOS) KAO CORP.

XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;

PI Okuda M, Saeki K;

XX WPI; 2002-437518/47.

XX New modified alkaline proteases useful in detergent compositions.

XX Claim 1; Page 10-11; 25pp; English.

XX This invention describes novel Bacillus sp. alkaline proteases useful in

CC detergent compositions, especially in laundry, bleaching or automatic

CC dishwasher detergents. The novel proteases have an increased detergency &

CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This

CC sequence represents a fragment of the alkaline protease KP43 from

CC Bacillus sp strain KSM-KP43 which is used to create the modified protease

CC represented in AAM50090

XX
 SQ

Sequence 434 AA;

Query Match 95.2%; Score 2143; DB 5; Length 434;

Best Local Similarity 93.5%; Pred. No. 1.5e-152;

Matches 406; Conservative 19; Mismatches 9; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDGRNDSSMHEAFRGKITAIYALGRTN 60

Db 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDGRNDSSMHEAFRGKITAIYALGRTN 60

QY 61 NNDPENGHGHVAGSVLNGTGNKGMAPQANLVFQSVMDNSGGLGLPSNVSTLFSQAYS 120

Db 61 NNDPENGHGHVAGSVLNGTGNKGMAPQANLVFQSVMDNSGGLGLPSNVSTLFSQAYS 120

QY 121 AGARIHTNSWGAPVNGAYTTDSRNVDYVVRKNDMAVLFAAGNEGPNNGGTISAPGTAKNAI 180

Db 121 AGARIHTNSWGAPVNGAYTTDSRNVDYVVRKNDMAVLFAAGNEGPNNGGTISAPGTAKNAI 180

QY 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTIFLSARSSLAPDSSF 240

Db 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTIFLSARSSLAPDSSF 240

QY 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKRGITPKPSLLKAALIAGATDIGLY 300

Db 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKRGITPKPSLLKAALIAGATDIGLY 300

QY 301 PSGNQGWGRVTLDKSLNVAFVNETSSLTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 360

Db 301 PSGNQGWGRVTLDKSLNVAFVNETSSLTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 360

QY 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGDGRNNVNFVINAPOSQGTYYVEVOAYN 420

Db 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGDGRNNVNFVINAPOSQGTYYVEVOAYN 420

QY 421 VPOGPQAFSLAIYN 434

Db 421 VPOGPQAFSLAIYN 434

RESULT 3

AAY17090

ID AAY17090 standard; protein; 640 AA.

XX AAY17090;

XX 20-MAR-2003 (revised)

DT 21-JUL-1999 (first entry)

XX Bacillus alkaline protease.

XX Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;

KW washing composition; oxidising agent.

XX Bacillus sp.

OS WO9918218-A1.

XX 15-APR-1999.

XX 07-OCT-1998; 98WO-JP004528.

XX 07-OCT-1997; 97JP-00274570.

XX (KAOS) KAO CORP.

XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;

PI Shikata S, Nomura M;

XX WPI; 1999-287736/27.

DR N-PSDB; AAX37278.

XX Alkali protease from Bacillus used in washing powders.

PI

XX Disclosure; Page 58-63; 71pp; Japanese.

XX The invention relates to alkaline proteases produced by strains of

XX Bacillus. The proteases ability to digest casein is not inhibited by

XX oleic acid and they have a high stability to oxidising agents. The

XX alkaline protease of the invention has the following properties: (a) it

XX is active over the pH range 4-13 and has at least 80% of its optimum

XX activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is

XX stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)

XX its ability to digest casein is not inhibited by oleic acid; (e) it has

XX molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be

XX used as enzymes in washing compositions for use in automatic dishwashers

XX and for washing clothes. The stability to oxidising agents allows the

XX enzyme to be an effective component of washing compositions including

XX bleaches. The present sequence represents an alkaline protease. (Updated

XX on 20-MAR-2003 to correct DR field.)

XX Sequence 640 AA;

SQ

Query Match 95.2%; Score 2143; DB 2; Length 640;

Best Local Similarity 93.5%; Pred. No. 2.5e-152;

Matches 406; Conservative 19; Mismatches 9; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITAIYALGRTN 60

Db 207 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITAIYALGRTN 266

QY 61 NANDPNHGHTHVAGSVLNGSTNKGMAPOANLVFQSVMDNSGGLGGLPSNVSTLFSQAYS 120

Db 267 NANDPNHGHTHVAGSVLNGSTNKGMAPOANLVFQSVMDNSGGLGGLPSNLQTLFSQAYS 326

QY 121 AGARIHTNSGAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNNGGTTISAPGTAKNAI 180

Db 327 AGARIHTNSGAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNNGGTTISAPGTAKNAI 386

QY 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTIFLSARSSLPDSSF 240

Db 387 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTIFLSARSSLPDSSF 446

QY 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFYKRGITPKPSLLKAALIGATDGLGY 300

Db 447 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFYKRGITPKPSLLKAALIGATDGLGY 506

QY 301 PSNGQWGRVTLDKSLNVAFVNETSLSLTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 360

Db 507 PSNGQWGRVTLDKSLNVAFVNETSLSLTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 566

QY 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWDRNENVENFINAPQSGTITVEQAYN 420

Db 567 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWDRNENVENFINAPQSGTITVEQAYN 626

QY 421 VPQGPQAFSLAIYN 434

Db 627 VPVGPQTFSLAIYN 640

RESULT 4

AAV17091

ID AAY17091 standard; protein; 640 AA.

XX

AC AAY17091;

XX

DT 20-MAR-2003 (revised)

DT 21-JUL-1999 (first entry)

XX

DE Bacillus alkaline protease.

XX

KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;

KW washing composition; oxidising agent.

XX

OS Bacillus sp.

XX

PN WO9918218-A1.

XX 15-APR-1999.

XX 07-OCT-1998; 98WO-JPC04528.

XX 07-OCT-1997; 97JP-00274570.

XX (KAOS) KAO CORP.

XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;

XX Shikata S, Nomura M;

XX WPI; 1999-287736/27.

DR N-PSDB; AAX37279.

XX Alkali protease from Bacillus used in washing powders.

XX Disclosure; Page 63-68; 71pp; Japanese.

XX The invention relates to alkaline proteases produced by strains of

XX Bacillus. The proteases ability to digest casein is not inhibited by

XX oleic acid and they have a high stability to oxidising agents. The

XX alkaline protease of the invention has the following properties: (a) it

XX is active over the pH range 4-13 and has at least 80% of its optimum

XX activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is

XX stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)

XX its ability to digest casein is not inhibited by oleic acid; (e) it has

XX molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be

XX used as enzymes in washing compositions for use in automatic dishwashers

XX and for washing clothes. The stability to oxidising agents allows the

XX enzyme to be an effective component of washing compositions including

XX bleaches. The present sequence represents an alkaline protease. (Updated

XX on 20-MAR-2003 to correct DR field.)

XX Sequence 640 AA;

SQ

Query Match 95.2%; Score 2141; DB 2; Length 640;

Best Local Similarity 93.5%; Pred. No. 3.6e-152;

Matches 406; Conservative 19; Mismatches 9; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITAIYALGRTN 60

Db 207 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITAIYALGRTN 266

QY 61 NANDPNHGHTHVAGSVLNGSTNKGMAPOANLVFQSVMDNSGGLGGLPSNVSTLFSQAYS 120

Db 267 NANDPNHGHTHVAGSVLNGSTNKGMAPOANLVFQSVMDNSGGLGGLPSNLQTLFSQAYS 326

QY 121 AGARIHTNSGAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNNGGTTISAPGTAKNAI 180

Db 327 AGARIHTNSGAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNNGGTTISAPGTAKNAI 386

QY 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTIFLSARSSLPDSSF 240

Db 387 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTIFLSARSSLPDSSF 446

QY 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFYKRGITPKPSLLKAALIGATDGLGY 300

Db 447 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFYKRGITPKPSLLKAALIGATDGLGY 506

QY 301 PSNGQWGRVTLDKSLNVAFVNETSLSLTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 360

Db 507 PSNGQWGRVTLDKSLNVAFVNETSLSLTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 566

QY 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWDRNENVENFINAPQSGTITVEQAYN 420

Db 567 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWDRNENVENFINAPQSGTITVEQAYN 626

QY 421 VPQGPQAFSLAIYN 434

Db 627 VPVGPQTFSLAIYN 640

```

RESULT 5
AAM50081
ID AAM50081 standard; protein; 434 AA.
XX
AC AAM50081;
XX
DT 12-AUG-2002 (first entry)
XX
DE Bacillus sp KSM-KP9860 alkaline protease protein fragment.
XX
KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX
OS Bacillus sp.
XX
PN EP1209233-A2.
XX
PD 29-MAY-2002.
XX
PF 22-NOV-2001; 2001EP-00127851.
XX
PR 22-NOV-2000; 2000JP-00355166.
XX
PR 12-APR-2001; 2001JP-00114048.
XX
PA (KAOS ) KAO CORP.
XX
PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
PI Okuda M, Saeki K;
XX
WPI; 2002-437518/47.
XX
New modified alkaline proteases useful in detergent compositions.
PT
PS Claim 5; Page 12-13; 25pp; English.
XX
This invention describes novel Bacillus sp. alkaline proteases useful in
detergent compositions, especially in laundry, bleaching or automatic
dishwasher detergents. The novel proteases have an increased detergency &
(34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
sequence represents a fragment of the alkaline protease KP9860 from
Bacillus sp strain KSM-KP9860 described in the method of the invention
XX
SQ Sequence 434 AA;
Query Match 94.7%; Score 2130; DB 5; Length 434;
Best Local Similarity 92.4%; Pred. No. 1.4e-151;
Matches 401; Conservative 24; Mismatches 9; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQSSYGLYGGQGVAVADTGLDGTGRNDSMHEAFRGKITAIYALGRTN 60
DB 1 NDVARGIVKADVAQSSYGLYGGQGVAVADTGLDGTGRNDSMHEAFRGKITAIYALGRTN 60
QY 61 NANDPENGHGHVAGSVLNGTSGKMAPQANLVFQSVMDNSGGLGLPSNVSTLPSOAYS 120
DB 61 NANDTNGHGHVAGSVLNGATNKMAPQANLVFQSIMDSSGGLGLPSNLQTLPSOAYS 120
QY 121 AGARIHTNSGAPVNGAYTTDSRVDYVRKNDMAVLFAAGNPGNGGTTISAPGTAKNAI 180
DB 121 AGARIHTNSGAAVNGAYTTDSRVDYVRKNDMTILFAAGNPGNGGTTISAPGTAKNAI 180
QY 181 TVGATENLRPSFGSYADINHNHVAQFSRGPTKGRIPKDPVMAPTFILGARSLLAPDSSE 240
DB 181 TVGATENLRPSFGSYADINHNHVAQFSRGPTKGRIPKDPVMAPTFILGARSLLAPDSSE 240
QY 241 WANHDSKYAYMGTSMTPIVAGNVAQLREHFINKRIGITEPKSLKALIAAGATDGLGY 300
DB 241 WANHDSKYAYMGTSMTPIVAGNVAQLREHFINKRIGITEPKSLKALIAAGATDGLGY 300
QY 301 PSNGCHGRVTLDKSLNVAFWNETSSLSNKTQKTSYFTAQSGKPLKISLVMSDAPASTA 360
DB 301 PNGCQGWGRVTLDKSLNVAFWNETSSLSNKTQKTSYFTATAGKPLKISLVMSDAPASTA 360
QY 361 SVTLVNDLVLITAPNGTRYGVNGDFSPAFDNDGNNVNFINSQSGTYTVEVQAYN 420

```

```

361 SVTLVNDLVLITAPNGTRYGVNGDFSPAFDNDGNNVNFINSQSGTYTVEVQAYN 420
QY 421 VPOGPQAFSLAIYN 434
DB 421 VPVGPNFSLAIYN 434
RESULT 6
AAY17089
ID AAY17089 standard; protein; 639 AA.
XX
AC AAY17089;
XX
DT 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX
DE Bacillus alkaline protease.
XX
KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
washing composition; oxidising agent.
XX
OS Bacillus sp.
XX
PN WO9918218-A1.
XX
PD 15-APR-1999.
XX
PF 07-OCT-1998; 98WO-JP004528.
XX
PR 07-OCT-1997; 97JP-00274570.
XX
PA (KAOS ) KAO CORP.
XX
PI Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;
XX
WPI; 1999-287736/27.
XX
N-PSDB; AAX37277.
PT Alkali protease from Bacillus used in washing powders.
XX
PS Disclosure; Page 53-58; 71pp; Japanese.
XX
The invention relates to alkaline proteases produced by strains of
Bacillus. The proteases ability to digest casein is not inhibited by
oleic acid and they have a high stability to oxidising agents. The
alkaline protease of the invention has the following properties: (a) it
is active over the pH range 4-13 and has at least 80% of its optimum
activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d)
its ability to digest casein is not inhibited by oleic acid; (e) it has
molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
used as enzymes in washing compositions for use in automatic dishwashers
and for washing clothes. The stability to oxidising agents allows the
enzyme to be an effective component of washing compositions including
bleaches. The present sequence represents an alkaline protease. (Updated
on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 639 AA;
Query Match 94.3%; Score 2122; DB 2; Length 639;
Best Local Similarity 92.2%; Pred. No. 9.5e-151;
Matches 400; Conservative 24; Mismatches 10; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQSSYGLYGGQGVAVADTGLDGTGRNDSMHEAFRGKITAIYALGRTN 60
DB 206 NDVARGIVKADVAQSSYGLYGGQGVAVADTGLDGTGRNDSMHEAFRGKITAIYALGRTN 265
QY 61 NANDPENGHGHVAGSVLNGTSGKMAPQANLVFQSVMDNSGGLGLPSNVSTLPSOAYS 120
DB 266 NANDTNGHGHVAGSVLNGATNKMAPQANLVFQSIMDSSGGLGLPSNLQTLPSOAYS 325

```

QY 121 AGARITNSWGPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
 DB 326 AGARITNSWGAANVGAYTTDSRNVDYVRKNDMTILFAAGNERPNGGTISAPGTAKNAI 385
 QY 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTIFLSARSSSLAPDSSF 240
 DB 386 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTIFLSARSSSLAPDSSF 445
 QY 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKRGITPKPSLLKAALIAAGADTIGLY 300
 DB 446 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKRGITPKPSLLKAALIAAGADTIGLY 505
 QY 301 PSGNQGWRVTLDKSLNVAFVNETSSLSTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 360
 DB 506 PNGNQGWGRVTLDKSLNVAFVNETSSLSTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 565
 QY 361 SVTLVNDLVLITAPNGTKYVGNDFAPYDNNWGRNENVENFINAPQSGTYTVEVQAYN 420
 DB 566 SVTLVNDLVLITAPNGTKYVGNDFAPYDNNWGRNENVENFINAPQSGTYTVEVQAYN 625
 QY 421 VPQGPQAFSLAIYN 434
 DB 626 VPVGPQNFSLAIYN 639

RESULT 7

AAW50086
 ID AAW50086 standard; protein; 433 AA.

AC AAW50086;

XX 12-AUG-2002 (first entry)

XX Bacillus sp alkaline protease protein A-2 fragment.

XX Alkaline protease; detergent; laundry; bleaching; dishwasher.

XX Bacillus sp.

XX EP1209233-A2.

XX 29-MAY-2002.

XX 22-NOV-2001; 2001EP-00127851.

XX 22-NOV-2000; 2000JP-00355166.

XX 12-APR-2001; 2001JP-00114048.

XX (KAOS) KAO CORP.

XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;

XX Okuda M, Saeki K;

XX WPI; 2002-437518/47.

XX New modified alkaline proteases useful in detergent compositions.

XX Claim 5; Page 20-21; 25pp; English.

XX This invention describes novel Bacillus sp. alkaline proteases useful in

XX detergent compositions, especially in laundry, bleaching or automatic

XX dishwasher detergents. The novel proteases have an increased detergency &

XX (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This

XX sequence represents a fragment of the alkaline protease A-2 from Bacillus

XX sp NCIB12513 described in the method of the invention

XX SQ Sequence 433 AA;

Query Match 94.1%; Score 2116.5; DB 5; Length 433;

Best Local Similarity 92.2%; Pred. No. 1.4e-150;

Matches 400; Conservative 21; Mismatches 12; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSVGLYGGQGVAVADTGLDGTGRNDSMHEAFRGKITAIYALGRTN 60

DB 1 NDVARGIVKADVAQSSVGLYGGQGVAVADTGLDGTGRNDSMHEAFRGKITAIYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLNGTGNKGMAPQANLVFQSVMDNSGGLGGLPSNVSTLFSQAYS 120
 DB 61 NANDPNHGHTHVAGSVLNGTGNKGMAPQANLVFQSVMDNSGGLGGLPSNVSTLFSQAYS 119
 QY 121 AGARIHTNSGAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
 DB 120 AGARIHTNSGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPNGGTISAPGTAKNAI 179
 QY 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTIFLSARSSSLAPDSSF 240
 DB 180 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTIFLSARSSSLAPDSSF 239
 QY 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKRGITPKPSLLKAALIAAGADTIGLY 300
 DB 240 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKRGITPKPSLLKAALIAAGADTIGLY 299
 QY 301 PSGNQGWRVTLDKSLNVAFVNETSSLSTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 360
 DB 300 PNGNQGWGRVTLDKSLNVAFVNETSSLSTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 359
 QY 361 SVTLVNDLVLITAPNGTKYVGNDFAPYDNNWGRNENVENFINAPQSGTYTVEVQAYN 420
 DB 360 SVTLVNDLVLITAPNGTKYVGNDFAPYDNNWGRNENVENFINAPQSGTYTVEVQAYN 419
 QY 421 VPQGPQAFSLAIYN 434
 DB 420 VPVGPQNFSLAIYN 433

RESULT 8

AAW89547

ID AAW89547 standard; protein; 641 AA.

AC AAW89547;

XX 12-APR-1999 (first entry)

XX Bacillus JP170 protease.

XX Protease; detergent; surfactant; leather processing; debittering;

XX flavour.

XX Bacillus sp.

XX Key Location/Qualifiers

XX Peptide 1..33

XX Region /note= "signal peptide"

XX Protein /note= "prepro region"

XX /note= "mature protein"

XX WO9856927-A2.

XX 17-DEC-1998.

XX 09-JUN-1998; 98WO-US012005.

XX 12-JUN-1997; 97US-00873479.

XX (NOVO) NOVO NORDISK BIOTECH INC.

XX Sloma A, Christianson L;

XX WPI; 1999-080908/07.

XX N-PSDB; AAW82382.

XX Novel protease from Bacillus subtilis LC20 - useful in laundry and

XX dishwashing detergents and for leather processing.

QY 421 VPOGQAFSLAIVN 434
 DB 626 VPVGQXFSLSAIVN 639

RESULT 10
 AAY17088
 ID AAY17088 standard; protein; 640 AA.
 XX AC AAY17088;
 XX 20-MAR-2003 (revised)
 DT 21-JUL-1999 (first entry)
 XX DE An alkaline protease sequence from Bacillus species.
 XX KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
 KW washing composition; oxidising agent.
 XX OS Bacillus sp.
 XX FT Key Location/Qualifiers
 FT Misc-difference 1.640
 FT /note= "all residues indicated as Xaa are arbitrary amino
 FT acids"
 XX WO9918218-A1.
 XX PD 15-APR-1999.
 XX PF 07-OCT-1998; 98WO-JP004528.
 XX PR 07-OCT-1997; 97JP-00274570.
 XX PA (KAOS) KAO CORP.
 XX PI Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
 PI Shikata S, Nomura M;
 XX WPI: 1999-287736/27.
 DR N-PSDB; AAX37278.
 XX PT Alkali protease from Bacillus used in washing powders.
 XX PS Claim 3; Page 50-53; 71pp; Japanese.
 XX CC The invention relates to alkaline proteases produced by strains of
 CC Bacillus. The proteases ability to digest casein is not inhibited by
 CC oleic acid and they have a high stability to oxidising agents. The
 CC alkaline protease of the invention has the following properties: (a) it
 CC is active over the pH range 4-13 and has at least 80% of its optimum
 CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
 CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has
 CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
 CC used as enzymes in washing compositions for use in automatic dishwashers
 CC and for washing clothes. The stability to oxidising agents allows the
 CC enzyme to be an effective component of washing compositions including
 CC bleaches. The present sequence represents an alkaline protease of the
 CC invention. (Updated on 20-MAR-2003 to correct DR field.)
 XX SQ Sequence 640 AA;

Query Match 92.5%; Score 2082; DB 2; Length 640;
 Best Local Similarity 91.5%; Fred. No. 9.7e-148;
 Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQSSVGLGQGVAVADTGLDGRNDSMHEAPRGKITAIYALGRN 60
 DB 207 NDVARGIVKADVAQSSVGLGQGVAVADTGLDGRNDSMHEAPRGKITAIYALGRN 266

QY 61 NANDPNHGHVAGSVLGNCTSNKGMAPQANLVFQSVMSNGGLGSLNSTLFSQAYS 120

DB 267 NANDTNGHGHVAGSVLGNCTSNKGMAPQANLVFQSVMSNGGLGSLNSTLFSQAYS 326
 QY 121 AGARIHTNSWGAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
 DB 327 AGARIHTNSWGAAVNGAYTTDSRNVDYVRKNDWTLFAAGNEXPNGGTISAPGTAKNAI 386
 QY 181 TVGATENLRPSFGSYANINHVAFQSSRGPTKGRIPKDVMAFGTFTLSARSLAPDSSF 240
 DB 387 TVGATENLRPSFGSYADNINHVAFQSSRGPTKGRIPKDVMAFGTFTLSARSLAPDSSF 446
 QY 241 WANHDSKVAYMGTSMTPIVAGNVAOLREHFVKNRGITPKPILLKALITAGATDGLGY 300
 DB 447 WANHDSKVAYMGTSMTPIVAGNVAOLREHFVKNRGITPKPILLKALITAGATDGLGY 506
 QY 301 PSGNQGWGRVTLDSLVNVPVNETSSLTNOKATYSPTAQSGKPLKISLVMSDAPASTSA 360
 DB 507 PNGNQGWGRVTLDSLVNVPVNETSSLTNOKATYSPTAQSGKPLKISLVMSDAPASTSA 566
 QY 361 SVTLVNDLVLVITAPNGTKVGVNDFTAPYDNNVDGRNVDNVPINAPQSGTYTVEVQAYN 420
 DB 567 SVTLVNDLVLVITAPNGTKVGVNDFTAPYDNNVDGRNVDNVPINAPQSGTYTVEVQAYN 626
 QY 421 VPOGQAFSLAIVN 434
 DB 627 VPVGQXFSLSAIVN 640

RESULT 11
 AAM50084
 ID AAM50084 standard; protein; 433 AA.
 XX AC AAM50084;
 XX DT 12-AUG-2002 (first entry)
 XX DE Bacillus sp SD-521 (FERM BP-11162) alkaline protease protein fragment.
 XX KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX OS Bacillus sp.
 XX FN EPI209233-A2.
 XX PD 29-MAY-2002.
 XX PF 22-NOV-2001; 2001EP-00127851.
 XX PR 22-NOV-2000; 2000JP-00355166.
 XX PR 12-APR-2001; 2001JP-00114048.
 XX PA (KAOS) KAO CORP.
 XX PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 XX Okuda M, Saeki K;
 XX WPI: 2002-437518/47.
 XX PT New modified alkaline proteases useful in detergent compositions.
 XX PS Claim 5; Page 16-18; 25pp; English.
 XX CC This invention describes novel Bacillus sp. alkaline proteases useful in
 XX detergent compositions, especially in laundry, bleaching or automatic
 XX dishwasher detergents. The novel proteases have an increased detergency &
 XX (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 XX sequence represents a fragment of the alkaline protease SD-521 from
 XX Bacillus sp strain SD-521 (FERM BP-11162) described in the method of the
 XX invention
 XX SQ Sequence 433 AA;

Query Match 90.0%; Score 2024.5; DB 5; Length 433;

Best Local Similarity 88.5%; Pred. NO. 1.2e-143;
Matches 384; Conservative 28; Mismatches 21; Indels 1; Gaps 1;
QY 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 60
DB 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 60
QY 61 NNDPNHGHTHVAGSVLNGTSGKGMAPQANLVFQSVMDNSGGLGGLPSNVSTLPSQAYS 120
DB 61 NNDPNHGHTHVAGSVLNGTSGKGMAPQANLVFQSVMDNSGGLGGLPSNVSTLPSQAYS 119
QY 121 AGARIHTNSWGAPVNGAYTTDSRNVDYVKKNDMAVLFAAGNEGPNNGGTISAPGTAKNAI 180
DB 120 AGARIHTNSWGAPVNGAYTTANSRQVDEYVYRNDMTVLFAAGNEGPNNGGTISAPGTAKNAI 179
QY 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPTGTFILSARSSLAPDSSF 240
DB 180 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPTGTFILSARSSLAPDSSF 239
QY 241 WANHSKYAYMGGTSMATPIVAGNVQALREHFIKRGITPKPSLLKAALITAGATDGLGY 300
DB 240 WANYSKYAYMGGTSMATPIVAGNVQALREHFIKRGITPKPSLLKAALITAGATDGLGY 299
QY 301 PSNGQGWGRVTLDKSLNVAFVNETSSLTNOKATYSFTAQSGKPLKISLVWSDAPASTSA 360
DB 300 PSNGQGWGRVTLDKSLNVAFVNETSSLTNOKATYSFTAQSGKPLKISLVWSDAPASTSA 359
QY 361 SVTLVNDLDELVTAPNGTKYVGNDFTPAYDNNWDRNNVNFVINAQSGTYYTVEQAYN 420
DB 360 SVTLVNDLDELVTAPNGTKYVGNDFTPAYDNNWDRNNVNFVINAQSGTYYTVEQAYN 419
QY 421 VPQGPQAFSLAIYN 434
DB 420 VPQGPQAFSLAIYH 433

RESULT 12
AAW50082
ID AAW50082 standard; protein; 433 AA.
XX AC AAW50082;
XX DT 12-AUG-2002 (first entry)
XX DE Bacillus sp D6- (FERM P1592) alkaline protease protein fragment.
XX KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX OS Bacillus sp.
XX PN EPI209233-A2.
XX PD 29-MAY-2002.
XX PF 22-NOV-2001; 2001EP-00127851.
XX PR 22-NOV-2000; 2000JP-00355166.
XX PR 12-APR-2001; 2001JP-00114048.
XX PA (KAO) KAO CORP.
XX PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
XX PI Okuda M, Saeki K;
XX DR WPI; 2002-437518/47.
XX PT New modified alkaline proteases useful in detergent compositions.
XX PS Claim 5; Page 13-15; 25pp; English.

This invention describes novel Bacillus sp. alkaline proteases useful in detergent compositions, especially in laundry, bleaching or automatic dishwasher detergents. The novel proteases have an increased detergency &

CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
CC sequence represents a fragment of the alkaline protease E-1 from Bacillus
CC sp strain D6- (FERM-P1592) described in the method of the invention
XX SQ Sequence 433 AA;
Query Match 89.8%; Score 2020.5; DB 5; Length 433;
Best Local Similarity 88.2%; Pred. No. 2.3e-143;
Matches 383; Conservative 28; Mismatches 22; Indels 1; Gaps 1;
QY 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 60
DB 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 60
QY 61 NNDPNHGHTHVAGSVLNGTSGKGMAPQANLVFQSVMDNSGGLGGLPSNVSTLPSQAYS 120
DB 61 NNDPNHGHTHVAGSVLNGTSGKGMAPQANLVFQSVMDNSGGLGGLPSNVSTLPSQAYS 119
QY 121 AGARIHTNSWGAPVNGAYTTDSRNVDYVKKNDMAVLFAAGNEGPNNGGTISAPGTAKNAI 180
DB 120 AGARIHTNSWGAPVNGAYTTANSRQVDEYVYRNDMTVLFAAGNEGPNNGGTISAPGTAKNAI 179
QY 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPTGTFILSARSSLAPDSSF 240
DB 180 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPTGTFILSARSSLAPDSSF 239
QY 241 WANHSKYAYMGGTSMATPIVAGNVQALREHFIKRGITPKPSLLKAALITAGATDGLGY 300
DB 240 WANYSKYAYMGGTSMATPIVAGNVQALREHFIKRGITPKPSLLKAALITAGATDGLGY 299
QY 301 PSNGQGWGRVTLDKSLNVAFVNETSSLTNOKATYSFTAQSGKPLKISLVWSDAPASTSA 360
DB 300 PSNGQGWGRVTLDKSLNVAFVNETSSLTNOKATYSFTAQSGKPLKISLVWSDAPASTSA 359
QY 361 SVTLVNDLDELVTAPNGTKYVGNDFTPAYDNNWDRNNVNFVINAQSGTYYTVEQAYN 420
DB 360 SVTLVNDLDELVTAPNGTKYVGNDFTPAYDNNWDRNNVNFVINAQSGTYYTVEQAYN 419
QY 421 VPQGPQAFSLAIYN 434
DB 420 VPQGPQAFSLAIYH 433

RESULT 13
AAW89548
ID AAW89548 standard; protein; 636 AA.
XX AC AAW89548;
XX DT 12-APR-1999 (first entry)
XX DE Bacillus sp. alkaline protease Y.
XX KW Alkaline protease Y; detergent; surfactant; leather processing;
XX KW debittering; flavour.
XX OS Bacillus sp.
XX PN WO9856927-A2.
XX PD 17-DEC-1998.
XX PF 09-JUN-1998; 98WO-US012005.
XX PR 12-JUN-1997; 97US-00873479.
XX PA (NOVO) NOVO NORDISK BIOTECH INC.
XX PI Sloma A, Christianson L;
XX DR WPI; 1999-080908/07.
XX PT Novel protease from Bacillus subtilis LC20 - useful in laundry and

PT dishwashing detergents and for leather processing.
 XX Claim 3; Page 55-56; 77pp; English.
 XX
 CC This is the amino acid sequence of a *Bacillus* sp. alkaline protease Y
 CC that is said to have good alkali and surfactant resistance and improved
 CC detergency. It shows 77% identity to a newly isolated protease (see
 CC AAR9547) of *Bacillus* sp. Jp170 (NCIB 12513). The invention provides
 CC vectors, recombinant host cells and methods for the recombinant
 CC production of such proteases. The protease are used in laundry and
 CC dishwashing detergents, for institutional and industrial cleaning, and
 CC for leather processing, as well as for debittering and enhancing the
 CC degree of hydrolysis of protein hydrolysates, for flavour development
 CC through hydrolysis of proteins, degradation of undesired peptides and in
 CC enzymatic synthesis of peptides. They have enhanced stability towards
 CC oxidation under alkaline conditions, e.g. towards bleaching agents of the
 CC peroxy type. The invention also provides mutant cells in which the
 CC protease activity is diminished. Such cells can be used for the
 CC production of heterologous recombinant proteins
 XX
 SQ Sequence 636 AA;
 Query Match 89.6%; Score 2016.5; DB 2; Length 636;
 Best Local Similarity 88.0%; Pred. No. 8e-143;
 Matches 382; Conservative 30; Mismatches 21; Indels 1; Gaps 1;
 QY 1 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDGRNDSMHEAFRGKITAIYALGRTN 60
 Db 204 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGRNDSMHEAFRGKITAIYALGRTN 263
 QY 61 NANDPNGHGHVAGSVLNGTSGKMAPOANLVFQSVMDNSGGLGGLPSNLTLFSQAWN 120
 Db 264 NASDPNGHGHVAGSVLGNAL-NKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWN 322
 QY 121 AGARIHTNSWGAPVNGAYTTDSRNDDYVRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
 Db 323 AGARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTLVFAAGNEGPNGGTISAPGTAKNAI 382
 QY 181 TVGATENLRPSFGSYADNINHVAQFSRGTQDGRKPDVMAQPGTIFLSARSSLAPDSF 240
 Db 383 TVGATENYRPSFGSIADNPNHIAQFSRGATDRGRIKPDVAPGTIFLSARSSLAPDSF 442
 QY 241 WANHSKYAYMGGTSMATPIVAGNVAQLREHFIKNGRITPKPSLLKAALIGATDGLGY 300
 Db 443 WANYSKYAYMGGTSMATPIVAGNVAQLREHFIKNGRITPKPSLLKAALIGATDGLGY 502
 QY 301 PSNGQGWGRVTLDKSLNVAFNVTSSLTNOKATYSFTAQSGKPLKISLVMSDAPASTA 360
 Db 503 PSQDQGWGRVTLDKSLNVAFNVTSSLTNOKATYSFTAQSGKPLKISLVMSDAPASTA 562
 QY 361 SVTLVNDLVLITAPNGTKYVGNDFTPYDNNWNGRNNVNFVINAPOSQGTYYVEVQAYN 420
 Db 563 SVTLVNDLVLITAPNGTKYVGNDFTPYDNNWNGRNNVNFVINAPOSQGTYYVEVQAYN 622
 QY 421 VPGQPQAFSLAIYN 434
 Db 623 VPSGPQRFSLAIHV 636
 RESULT 14
 AAR26274
 ID AAR26274 standard; protein; 433 AA.
 AC AAR26274;
 XX
 XX 05-FEB-1993 (first entry)
 DT
 XX Alkali-protease Ya enzyme.
 DE
 XX Alkali resistance; surface active agent resistance; detergency improver.
 KW
 OS *Bacillus* sp. Y.
 XX

PN JF04197182-A.
 XX
 PD 16-JUL-1992.
 XX
 XX 28-NOV-1990; 90JP-00327110.
 PF
 XX 28-NOV-1990; 90JP-00327110.
 PR
 XX (LLOY) LION CORP.
 PA
 XX WPI; 1992-288440/35.
 DR N-PSDB; AAQ27516.
 DR
 XX DNA coding alkali-protease Ya enzyme - has good alkali and surfactant
 PT resistance and improves detergency.
 PT
 XX Claim 2; Page 1; 17pp; Japanese.
 PS
 CC The sequence is that of alkali-protease Ya enzyme which can be used in
 CC the recombinant production of Ya enzyme. Ya enzyme is excellent in alkali
 CC resistance and surface active agent resistance and improves detergency
 CC
 XX SQ Sequence 433 AA;
 Query Match 89.4%; Score 2010.5; DB 2; Length 433;
 Best Local Similarity 87.8%; Pred. No. 1.3e-142;
 Matches 381; Conservative 30; Mismatches 22; Indels 1; Gaps 1;
 QY 1 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDGRNDSMHEAFRGKITAIYALGRTN 60
 Db 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGRNDSMHEAFRGKITAIYALGRTN 60
 QY 61 NANDPNGHGHVAGSVLNGTSGKMAPOANLVFQSVMDNSGGLGGLPSNLTLFSQAYS 120
 Db 61 NASDPNGHGHVAGSVLGNAL-NKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWN 119
 QY 121 AGARIHTNSWGAPVNGAYTTDSRNDDYVRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
 Db 120 AGARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTLVFAAGNEGPNGGTISAPGTAKNAI 179
 QY 181 TVGATENLRPSFGSYADNINHVAQFSRGTQDGRKPDVMAQPGTIFLSARSSLAPDSF 240
 Db 180 TVGATENYRPSFGSIADNPNHIAQFSRGATDRGRIKPDVAPGTIFLSARSSLAPDSF 239
 QY 241 WANHSKYAYMGGTSMATPIVAGNVAQLREHFIKNGRITPKPSLLKAALIGATDGLGY 300
 Db 240 WANYSKYAYMGGTSMATPIVAGNVAQLREHFIKNGRITPKPSLLKAALIGATDGLGY 299
 QY 301 PSNGQGWGRVTLDKSLNVAFNVTSSLTNOKATYSFTAQSGKPLKISLVMSDAPASTA 360
 Db 300 PSQDQGWGRVTLDKSLNVAFNVTSSLTNOKATYSFTAQSGKPLKISLVMSDAPASTA 359
 QY 361 SVTLVNDLVLITAPNGTKYVGNDFTPYDNNWNGRNNVNFVINAPOSQGTYYVEVQAYN 420
 Db 360 SVTLVNDLVLITAPNGTKYVGNDFTPYDNNWNGRNNVNFVINAPOSQGTYYVEVQAYN 419
 QY 421 VPGQPQAFSLAIYN 434
 Db 420 VPSGPQRFSLAIHV 433
 RESULT 15
 AAR61495
 ID AAR61495 standard; protein; 433 AA.
 AC AAR61495;
 XX
 XX 06-NOV-1998 (first entry)
 DT
 XX Modified *Bacillus* lion Y protease.
 DE
 XX *Bacillus* lion Y protease; polyethylene glycol; PEG; soap;
 KW methoxypolyethyleneglycol; mPEG; skin; hair care product; cosmetic;
 XX

KW lipstick; hair gel; sun oil; shampoo; hair dye; insect repellent.
 XX Bacillus sp.
 XX Key Location/Qualifiers
 XX Modified-site 1..433
 FT /note= "The enzyme is modified by methoxypolyethyleneglycol
 FT molecules covalently attached to the N-terminal amino
 FT group and to fourteen unspecified amino groups of lysine
 FT residues present on the surface of the enzyme"
 XX
 PN WO9830682-A1.
 XX
 XX 16-JUL-1998.
 XX
 XX 12-JAN-1998; 98WO-DK000015.
 XX
 XX 10-JAN-1997; 97DK-00000038.
 PR 25-JUN-1997; 97DK-00000754.
 XX
 XX (NOVO) NOVO-NORDISK AS.
 XX
 XX Olsen AA, Prento A;
 XX WPI; 1998-399132/34.
 XX
 XX New enzyme modified by attachment of many polymeric molecules - useful in
 PT skin and hair care products, has reduced tendency to cause sensitisation
 PT and increased stability.
 XX
 XX Claim 16; Page 44-45; 56pp; English.
 XX
 CC The present sequence represents the Bacillus lion Y protease. The
 CC invention claims for enzymes covalently modified on their surface by the
 CC attachment of a large number of small polymeric molecules, e.g.
 CC polyethylene glycols (PEG). The polymeric molecules are coupled to the N-
 CC terminal amino group and the amino groups of lysine residues found on the
 CC surface of the enzyme. In the example given, the lion Y protease was
 CC modified using methoxypolyethyleneglycol (mPEG) as the polymeric molecule.
 CC The N-terminal amino group and the amino groups of the fourteen lysine
 CC residues present on the surface of the lion Y protease were modified.
 CC Modification of the enzymes increases the stability and/or reduces the
 CC sensitising potential (allergenicity) of the enzyme, without
 CC significantly reducing enzymatic activity. Also, using a large number of
 CC relatively small polymeric molecules, rather than a few very large ones,
 CC provides a more even effect with reduced activity loss. The modified
 CC enzymes are claimed to be useful as components of a wide range of skin
 CC and hair care products, e.g. soaps, cosmetics, creams, gels, lipsticks,
 CC hair gels, sun oils, shampoos, hair dyes, insect repellants, etc
 XX
 SQ Sequence 433 AA;

Query Match 89.4%; Score 2010.5; DB 2; Length 433;
 Best Local Similarity 87.8%; Pred. No. 1.3e-142;
 Matches 381; Conservative 30; Mismatches 22; Indels 1; Gaps 1;
 QY 1 NDVARGIVKADVAQSSVGLYGGQGVAVADTGLDTCGRNDSMHEAPRGKITAIYALGRTN 60
 DB 1 NDVARGIVKADVAQNNYGLYGGQLVAVADTGLDTCGRNDSMHEAPRGKITAIYALGRTN 50
 QY 61 NANDPNHGHTHVAGSVLGNSTSKGMAPQANLVFQVMDNSGGLGGLPSNVSTLFSQAYS 120
 DB 61 NASDPNGHGHVAGSVLGNAL-NKGMAPQANLVFQSIMDSGGLGGLPSNLTLFSQAWN 119
 QY 121 AGARIHTNSGAPFVNGAYTTDSRVDDYVRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
 DB 120 AGARIHTNSGAPFVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAI 179
 QY 181 TVGATENLRPSFGSADNINHVAQFSRGPTKGRIKPDVMAFGTTILSARSLAPDSSF 240
 DB 180 TVGATENLRPSFGSIADPNHIAQFSRGATRGRIKPDVADPTTILSARSLAPDSSF 239
 QY 241 WANHDSKYAYMGTSNATPIVAGNVQLRHFIFKNRGIITPKPSLLKAAIAGATDIGLY 300

DB 240 WAINYSKIAYMGTSNATPIVAGNVQLRHFIFKNRGIITPKPSLLKAAIAGATDIGLY 299
 QY 301 PSNGQGWGRVTLDKSLNVAFVNETSSLSTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 360
 DB 300 PSNGQGWGRVTLDKSLNVAFVNETSSLSTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 359
 QY 361 SVTLVNDLDELIVITAPNGTKYVGNDFTPAYDNNWDGNNVNFINAPQSGTYTVEQAYN 420
 DB 360 SYTLVNDLDELIVITAPNGTKYVGNDFTPAYDNNWDGNNVNFINAPQSGTYTVEQAYN 419
 QY 421 VPOGPQAFSLAIYN 434
 DB 420 VPSGPQRFSLAIHV 433

Search completed: March 31, 2004, 16:04:33
 Job time : 48.2392 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:59:39 ; Search time 14.0215 Seconds
(without alignments)
1597.947 Million cell updates/sec

Title: US-09-985-689A-6
Perfect score: 2250
Sequence: 1 NDVARGIVKADVAQSSYGLY.....EVQAYNPQGPQAPSLAIVN 434

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/pCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------------------------|
| 1 | 2143 | 95.2 | 640 | 4 | US-09-509-814A-6 Sequence 6, Appli |
| 2 | 2141 | 95.2 | 640 | 4 | US-09-509-814A-8 Sequence 8, Appli |
| 3 | 2122 | 94.3 | 639 | 4 | US-09-509-814A-4 Sequence 4, Appli |
| 4 | 2116.5 | 94.1 | 641 | 2 | US-08-873-479-42 Sequence 42, Appli |
| 5 | 2082 | 92.5 | 639 | 4 | US-09-509-814A-1 Sequence 1, Appli |
| 6 | 2082 | 92.5 | 640 | 4 | US-09-509-814A-2 Sequence 2, Appli |
| 7 | 2010.5 | 89.4 | 433 | 4 | US-09-104-622A-4 Sequence 4, Appli |
| 8 | 2010.5 | 89.4 | 433 | 4 | US-09-019-532-4 Sequence 4, Appli |
| 9 | 2010.5 | 89.4 | 433 | 4 | US-09-338-746-4 Sequence 4, Appli |
| 10 | 2010.5 | 89.4 | 635 | 2 | US-08-873-479-43 Sequence 43, Appli |
| 11 | 1562.5 | 69.4 | 345 | 4 | US-09-512-251A-10 Sequence 10, Appli |
| 12 | 1562.5 | 69.4 | 345 | 4 | US-09-515-150A-10 Sequence 10, Appli |
| 13 | 1562.5 | 69.4 | 345 | 4 | US-09-281-13 Sequence 13, Appli |
| 14 | 450.5 | 20.0 | 659 | 3 | US-08-894-818B-1 Sequence 1, Appli |
| 15 | 450.5 | 20.0 | 659 | 4 | US-09-445-472-12 Sequence 12, Appli |
| 16 | 404 | 18.0 | 412 | 4 | US-09-445-472-1 Sequence 1, Appli |
| 17 | 404 | 18.0 | 522 | 3 | US-08-894-818B-3 Sequence 3, Appli |
| 18 | 404 | 18.0 | 522 | 4 | US-09-445-472-4 Sequence 4, Appli |
| 19 | 404 | 18.0 | 654 | 3 | US-08-894-818B-35 Sequence 35, Appli |
| 20 | 404 | 18.0 | 654 | 3 | US-09-445-472-16 Sequence 16, Appli |
| 21 | 403 | 17.9 | 659 | 3 | US-08-894-818B-5 Sequence 5, Appli |
| 22 | 355 | 15.8 | 520 | 3 | US-09-000-016-7 Sequence 7, Appli |
| 23 | 355 | 15.8 | 734 | 3 | US-09-514-340-7 Sequence 4, Appli |
| 24 | 355 | 15.8 | 734 | 3 | US-09-000-016-4 Sequence 4, Appli |
| 25 | 355 | 15.8 | 734 | 3 | US-09-514-340-4 Sequence 2, Appli |
| 26 | 355 | 15.8 | 823 | 3 | US-09-000-016-2 Sequence 2, Appli |
| 27 | 355 | 15.8 | 823 | 4 | US-09-514-340-2 Sequence 2, Appli |

| | | | | | | |
|----|-------|------|------|---|-------------------|--------------------|
| 28 | 316.5 | 14.1 | 903 | 1 | US-08-750-532-1 | Sequence 1, Appli |
| 29 | 316.5 | 14.1 | 1398 | 1 | US-08-750-532-9 | Sequence 9, Appli |
| 30 | 316.5 | 14.1 | 1398 | 3 | US-08-894-818B-8 | Sequence 8, Appli |
| 31 | 316.5 | 14.1 | 1398 | 4 | US-09-445-472-6 | Sequence 6, Appli |
| 32 | 290 | 12.9 | 237 | 1 | US-08-750-532-18 | Sequence 18, Appli |
| 33 | 274.5 | 12.2 | 418 | 2 | US-08-873-479-44 | Sequence 44, Appli |
| 34 | 272 | 12.1 | 418 | 4 | US-09-966-921A-2 | Sequence 2, Appli |
| 35 | 267 | 11.9 | 269 | 1 | US-07-706-691G-4 | Sequence 4, Appli |
| 36 | 267 | 11.9 | 269 | 2 | US-08-254-021-4 | Sequence 4, Appli |
| 37 | 267 | 11.9 | 269 | 2 | US-08-618-446-4 | Sequence 4, Appli |
| 38 | 267 | 11.9 | 269 | 3 | US-08-980-135-4 | Sequence 4, Appli |
| 39 | 267 | 11.9 | 269 | 4 | US-09-585-798-4 | Sequence 4, Appli |
| 40 | 266 | 11.8 | 269 | 1 | US-08-566-369-11 | Sequence 11, Appli |
| 41 | 266 | 11.8 | 269 | 3 | US-09-074-331-11 | Sequence 11, Appli |
| 42 | 266 | 11.8 | 269 | 5 | PCT-US95-01937-11 | Sequence 11, Appli |
| 43 | 265 | 11.8 | 269 | 1 | US-07-706-691G-5 | Sequence 5, Appli |
| 44 | 265 | 11.8 | 269 | 1 | US-08-254-021-5 | Sequence 5, Appli |
| 45 | 265 | 11.8 | 269 | 2 | US-08-618-446-5 | Sequence 5, Appli |

ALIGNMENTS

RESULT 1
US-09-509-814A-6
; Sequence 6, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 6
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-6

| | | | | | |
|-----------------------|-----|---|---------------------|-----------|-------------|
| Query Match | | 95.2% | Score 2143; | DB 4; | Length 640; |
| Best Local Similarity | | 93.5% | Pred. No. 2.7e-165; | Indels 0; | Gaps 0; |
| Matches 406; | | Conservative 19; | Mismatches 9; | | |
| Qy | 1 | NDVARGIVKADVAQSSYGLYGGQVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN | 60 | | |
| Db | 207 | NDVARGIVKADVAQSSYGLYGGQVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN | 266 | | |
| Qy | 61 | NANDPNGHGHVAGSVLNGTSGNKWAPQANLVFQSVMDSGGLGLPSNVSTLFSQAYS | 120 | | |
| Db | 267 | NANDTNGHGHVAGSVLNGTSGNKWAPQANLVFQSVMDSGGLGLPSNLTLFSQAYS | 326 | | |
| Qy | 121 | AGARIHNSNGAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNNGTTISAPGTAKNAI | 180 | | |
| Db | 327 | AGARIHNSNGAAVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPNNGTTISAPGTAKNAI | 386 | | |
| Qy | 181 | TVGATENLRPSFGSYADNINHVAFSSRGPTKGRKPDVWAGPTFLSARSLAPDSSF | 240 | | |
| Db | 387 | TVGATENLRPSFGSYADNINHVAFSSRGPTKGRKPDVWAGPTFLSARSLAPDSSF | 446 | | |
| Qy | 241 | WANHDSKYVMGTSMTPIVAGNVAQLREHFKIKRGITPKPSILKAAIAGATDGLGY | 300 | | |

Db 447 WANHDSKYAMGTSMTATPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAADIGLGY 506
Qy 301 PSNGQGWGRTLDKSLNVAFVNETSSLTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 360
Db 507 PNGNQGWGRTLDKSLNVAFVNETSSLTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 566
Qy 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNDGGRNNVENFINAPQSGTYTIEVQAYN 420
Db 567 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNDGGRNNVENFINAPQSGTYTIEVQAYN 626
Qy 421 VPOGPOAFSLAIYN 434
Db 627 VPVGPOFSLAIYN 640
RESULT 2
US-09-509-814A-8
; Sequence 8, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUW
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-OPCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-8
Query Match 95.2%; Score 2141; DB 4; Length 640;
Best Local Similarity 93.5%; Pred. No. 3.9e-165;
Matches 406; Conservative 19; Mismatches 9; Indels 0; Gaps 0;
Qy 1 NDVARGIVKADVAGSSVGLYGQGVAVADTGLDTGRNDSMHEAFRGKITAIYALGRTN 60
Db 207 NDVARGIVKADVAGSSVGLYGQGVAVADTGLDTGRNDSMHEAFRGKITAIYALGRTN 266
Qy 61 NANDPNHGHTHAGSVLNGTSTNKGMAPOANLVFQSVMDNSGGLGPLSNVSTLFSQAYS 120
Db 267 NANDTNGHGHVAGSVLNGSTNKGMAPOANLVFQSVMDNSGGLGPLSNVSTLFSQAYS 326
Qy 121 AGARIHTNSGAPVNGAYTTDSNRVDDYVRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
Db 327 AGARIHTNSGAAVNGAYTTDSNRVDDYVRKNDMTILFAAGNEGPNGGTISAPGTAKNAI 386
Qy 181 TVGATENLRPSFGSYADNINHAQFSRSGPTKDGRIKPDVMAFGTPILSARSSLAPDSSF 240
Db 387 TVGATENLRPSFGSYADNINHAQFSRSGPTKDGRIKPDVMAFGTPILSARSSLAPDSSF 446
Qy 241 WANHDSKYAMGTSMTATPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAADIGLGY 300
Db 447 WANHDSKYAMGTSMTATPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAADIGLGY 506
Qy 301 PSNGQGWGRTLDKSLNVAFVNETSSLTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 360
Db 507 PNGNQGWGRTLDKSLNVAFVNETSSLTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 566

Qy 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNDGGRNNVENFINAPQSGTYTIEVQAYN 420
Db 567 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNDGGRNNVENFINAPQSGTYTIEVQAYN 626
Qy 421 VPOGPOAFSLAIYN 434
Db 627 VPVGPOFSLAIYN 640
RESULT 3
US-09-509-814A-4
; Sequence 4, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUW
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-OPCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-4
Query Match 94.3%; Score 2122; DB 4; Length 639;
Best Local Similarity 92.2%; Pred. No. 1.4e-163;
Matches 400; Conservative 24; Mismatches 10; Indels 0; Gaps 0;
Qy 1 NDVARGIVKADVAGSSVGLYGQGVAVADTGLDTGRNDSMHEAFRGKITAIYALGRTN 60
Db 206 NDVARGIVKADVAGSSVGLYGQGVAVADTGLDTGRNDSMHEAFRGKITAIYALGRTN 265
Qy 61 NANDPNHGHTHAGSVLNGTSTNKGMAPOANLVFQSVMDNSGGLGPLSNVSTLFSQAYS 120
Db 266 NANDTNGHGHVAGSVLNGATNKGMAPOANLVFQSVMDNSGGLGPLSNVSTLFSQAYS 325
Qy 121 AGARIHTNSGAPVNGAYTTDSNRVDDYVRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
Db 326 AGARIHTNSGAAVNGAYTTDSNRVDDYVRKNDMTILFAAGNEGPNGGTISAPGTAKNAI 385
Qy 181 TVGATENLRPSFGSYADNINHAQFSRSGPTKDGRIKPDVMAFGTPILSARSSLAPDSSF 240
Db 386 TVGATENLRPSFGSYADNINHAQFSRSGPTKDGRIKPDVMAFGTPILSARSSLAPDSSF 445
Qy 241 WANHDSKYAMGTSMTATPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAADIGLGY 300
Db 446 WANHDSKYAMGTSMTATPIVAGNVAQLREHFVKNRGITPKPSLLKAAALIAAADIGLGY 505
Qy 301 PSNGQGWGRTLDKSLNVAFVNETSSLTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 360
Db 506 PNGNQGWGRTLDKSLNVAFVNETSSLTNQKATYSFTAQSGKPLKISLVMSDAPASTSA 565
Qy 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNDGGRNNVENFINAPQSGTYTIEVQAYN 420
Db 566 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNDGGRNNVENFINAPQSGTYTIEVQAYN 625
Qy 421 VPOGPOAFSLAIYN 434
Db 626 VPVGPOFSLAIYN 639

```

RESULT 4
US-08-873-479-42
; Sequence 42, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Lynne, Christianson
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSES: NO. 58917010 No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agris, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
; US-08-873-479-42

Query Match          94.1%; Score 2116.5; DB 2; Length 641;
Best Local Similarity 92.2%; Pred. No. 3.8e-163;
Matches 400; Conservative 21; Mismatches 12; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDTGRNDSMHEAPRGKITAIYALGRTN 60
Db 209 NDVARGIVKADVAQNNGLYGGQGVAVADTGLDTGRNDSMHEAPRGKITAIYALGRTN 268
QY 61 NANDPNGHGTHVAGSVLNGTSSNKGMAPQANLVFQSVYMDNSGGLGLPNSVTLFSQAYS 120
Db 269 NANDPNGHGTHVAGSVLNGAT-NKGMAPQANLVFQSIMDSGGGLGLPANLQTLFSQAYS 327
QY 121 AGARIHTNSWGAPVNGAYTTDSRNVDVYRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
Db 328 AGARIHTNSWGAPVNGAYTTDSRNVDVYRKNDMTILFAAGNEGPGSGGTISAPGTAKNAI 387
QY 181 TVGATENLRPSFGSADNINHVAQPSRRGPTKDGRIKPDVMAFCTFILSARSSLAPDSGF 240
Db 388 TVGATENLRPSFGSADNINHVAQPSRRGPTRDGRIKPDVMAFGTIVLSARSSLAPDSGF 447
QY 241 WANHDSKYAYMGSTWATPIVAGNVAOLREHFINKRGITPKBSLLKAALTAGATDIGLY 300
Db 448 WANHDSKYAYMGSTWATPIVAGNVAQUREHFVNRNGVTPKPSLLKAALTAGADVGLGF 507
QY 301 PSNGQGWGRVTLDKSLNVAFVNNETSSLTNOKATYSFTAQSGKPLKISLVWSDAPASTA 360
Db 508 PNCNGQGWGRVTLDKSLNVAFVNNETSPLTSOKATYSFTAQSGKPLKISLVWSDAPGSTA 567

```

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid

NAME/KEY: misc feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1

Query Match 92.5%; Score 2082; DB 4; Length 639;
Best Local Similarity 91.5%; Pred. No. 2.4e-160; Indels 0; Gaps 0;
Matches 397; Conservative 14; Mismatches 23;
QY 1 NDVARGIVKADVAQSSYGLYGGQGVAVADTGLDGRNDSMHEAFPGKITAIYALGRTN 60
Db |||||
206 NDVARGIVKADVAQSSYGLYGGQGVAVADTGLDGRNDSMHEAFPGKITAIYALGRTN 265
QY 61 NANDPNGHGHVAGSVLNGTSGKMAPQANLVFQSVMDNSGGLGGLPSNVSTLFSOAYS 120
Db |||||
266 NANDPNGHGHVAGSVLNGTSGKMAPQANLVFQSVMDNSGGLGGLPSNLQTLFSOAXS 325
QY 121 AGARIHTNSMGAPVNGAYTTDSRVNDDYVRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
Db |||||
326 AGARIHTNSMGAAVNGAYTTDSRVNDDYVRKNDMTILFAAGNEKPNGGTISAPGTAKNAI 385
QY 181 TVGATENLRPSFGSYADNINHVAFQSSRGPTKGRIPKDVMACTFTLSARSLAPDSSF 240
Db |||||
386 TVGATENLRPSFGSYADNINHVAFQSSRGPTKGRIPKDVMACTFTLSARSLAPDSSF 445
QY 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFIKNRGITPKPSLLKAALTAGATDGLGY 300
Db |||||
446 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFIKNRGITPKPSLLKAALTAGAAXGLGY 505
QY 301 PSGNQGGRVTLDKSLNVAFVNETSSLTQKATYSTAQSGPKLKSILVWSDAPASTSA 360
Db |||||
506 PNGNQGGRVTLDKSLNVAYVYNESSXLSLTSQKATYFTATAGKPLKISLVWSDAPASTTA 565
QY 361 SVTLVNDLVLVITAPNGTKYVGNDFTPAYDNNMWDGRNNVNFVFNAPQSGTYTVEVOAYN 420
Db |||||
566 SVTLVNDLVLVITAPNGTKYVGNDFTPAYDNNMWDGRNNVNFVFNAPQSGTYTVEVOAYN 625
QY 421 VPQGPQAPSLAIVN 434
Db |||||
626 VPVGPQXPFLAIVN 639

RESULT 6
US-09-509-814A-2
Sequence 2, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: TAKAIWA, MIKIO
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUW
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-OPCT
CURRENT APPLICATION NUMBER: US/09/509,814A

;; CURRENT FILING DATE: 2000-04-06
;; PRIOR APPLICATION NUMBER: PCI/JP98/04528
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: JP 9-274570
;; PRIOR FILING DATE: 1997-06-08
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 2
;; LENGTH: 640
;; TYPE: PRT
;; ORGANISM: Bacillus sp.
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (3)..(3)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (24)..(24)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (30)..(30)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (33)..(33)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (47)..(47)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (48)..(48)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (54)..(54)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (71)..(71)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (75)..(75)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (90)..(90)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (103)..(103)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (106)..(106)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (129)..(129)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (131)..(131)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (132)..(132)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (133)..(133)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (134)..(134)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (147)..(147)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (149)..(149)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (161)..(161)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature

;; LOCATION: (166)..(166)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (173)..(173)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (184)..(184)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (188)..(188)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (189)..(189)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (190)..(190)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (195)..(195)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (287)..(287)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (307)..(307)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (325)..(325)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (370)..(370)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (432)..(432)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (502)..(502)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (532)..(532)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (542)..(542)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (585)..(585)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (592)..(592)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (593)..(593)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (595)..(595)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (596)..(596)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (597)..(597)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (612)..(612)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (633)..(633)
;; OTHER INFORMATION: Xaa is any amino acid
;; OTHER INFORMATION: Xaa is any amino acid

US-09-509-814A-2

Query Match 92.5%; Score 2082; DB 4; Length 640;
Best Local Similarity 91.5%; Pred. No. 2.4e-160;
Matches 397; Conservative 14; Mismatches 23; Indels 0; Gaps 0;

```
QY 1 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDGRNDSSMHBAFRGKITAIYALGRTN 60
Db 207 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDGRNDSSMHBAFRGKITAIYALGRTN 266
QY 61 NANDPNHGHTHVAGSVLNGTSGKGMAPQANLVFQSVMDSSNGGLGGLPSNVSTLFSQAYS 120
Db 267 NANDPNHGHTHVAGSVLNGTSGKGMAPQANLVFQSVMDSSNGGLGGLPSNVSTLFSQAYS 326
QY 121 AGARIHTNSGAPVNGAYTTDSRVDDYVRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
Db 327 AGARIHTNSGAPVNGAYTTDSRVDDYVRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 386
QY 181 TVGATENLRPSFGSYADNINHVAQFSRGTGDKRIKPDVMAFGTFLSARSLAPDSSF 240
Db 387 TVGATENLRPSFGSYADNINHVAQFSRGTGDKRIKPDVMAFGTFLSARSLAPDSSF 446
QY 241 WANHDSKIAYMGGTSMATPIVAGNVQALREHFIKNGRITPKPSLLKAALIAGATDGLGY 300
Db 447 WANHDSKIAYMGGTSMATPIVAGNVQALREHFIKNGRITPKPSLLKAALIAGATDGLGY 506
QY 301 PSNGQGWGRVTLDKSLNVAFVNETSSLTNOKATYSFTAQSGKPLKISLVWSDAPASTA 360
Db 507 PSNGQGWGRVTLDKSLNVAFVNETSSLTNOKATYSFTAQSGKPLKISLVWSDAPASTA 566
QY 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGRNNVNFVINAPOSQGTYYVEQAYN 420
Db 567 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGRNNVNFVINAPOSQGTYYVEQAYN 626
QY 421 VPOGPOAFSLAIYN 434
Db 627 VPOGPOAFSLAIYN 640
```

RESULT 7

```
US-09-104-623A-4
; Sequence 4, Application US/09104623A
; Patent No. 6303752
; GENERAL INFORMATION:
; APPLICANT: Olsen, Arne Agerlin
; APPLICANT: Fatum, Tine Muxoll
; APPLICANT: Deussen, Heinz-Josef
; APPLICANT: Roggen, Erwin Ludc
; TITLE OF INVENTION: A Modified Polypeptide
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6303752o No. 6303752disk of No. 6303752th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,623A
; FILING DATE: 25-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 5256.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids
; TYPE: amino acid
; STRANDEDNESS:
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; STRAIN: Bacillus sp. Y
US-09-104-623A-4
```

```
Query Match 89.4%; Score 2010.5; DB 4; Length 433;
Best Local Similarity 87.8%; Pred. No. 8.4e-155;
Matches 381; Conservative 30; Mismatches 22; Indels 1; Gaps 1;
```

```
QY 1 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDGRNDSSMHBAFRGKITAIYALGRTN 60
Db 1 NDVARGIVKADVAQNNYGLYGQQLVAVADTGLDGRNDSSMHBAFRGKITAIYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLNGTSGKGMAPQANLVFQSVMDSSNGGLGGLPSNVSTLFSQAYS 120
Db 61 NANDPNHGHTHVAGSVLNGAL-NKGMAPQANLVFQSIMDSSGGLGGLPSNLTLSQAWN 119
QY 121 AGARIHTNSGAPVNGAYTTDSRVDDYVRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
Db 120 AGARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTLVFAAGNEGPNSGTISAPGTAKNAI 179
QY 181 TVGATENLRPSFGSYADNINHVAQFSRGTGDKRIKPDVMAFGTFLSARSLAPDSSF 240
Db 180 TVGATENLRPSFGSIADNPNHIAQFSRGTGDKRIKPDVMAFGTFLSARSLAPDSSF 239
QY 241 WANHDSKIAYMGGTSMATPIVAGNVQALREHFIKNGRITPKPSLLKAALIAGATDGLGY 300
Db 240 WANYSKIAYMGGTSMATPIVAGNVQALREHFIKNGRITPKPSLLKAALIAGATDGLGY 299
QY 301 PSNGQGWGRVTLDKSLNVAFVNETSSLTNOKATYSFTAQSGKPLKISLVWSDAPASTA 360
Db 300 PSNGQGWGRVTLDKSLNVAFVNETATATGCKATYSFTAQSGKPLKISLVWDAPGSTTA 359
QY 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGRNNVNFVINAPOSQGTYYVEQAYN 420
Db 360 SYTLVNDLVLITAPNGQKYVGNDFSYDNNWGRNNVNFVINAPOSQGTYYVEQAYN 419
QY 421 VPOGPOAFSLAIYN 434
Db 420 VPSGPQRFSLAIHV 433
```

RESULT 8

```
US-09-019-532-4
; Sequence 4, Application US/09019532B
; Patent No. 6416756
; GENERAL INFORMATION:
; APPLICANT: Olsen, Arne Agerlin
; APPLICANT: Prent, Amette
; TITLE OF INVENTION: A Modified Enzyme for Skin Care
; FILE REFERENCE: 4922.204-US
; CURRENT APPLICATION NUMBER: US/09/019,532B
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: 0038/97
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 0754/97
; EARLIER FILING DATE: 1997-06-25
; EARLIER APPLICATION NUMBER: 60/051,381
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: PCT/DK98/00015
; EARLIER FILING DATE: 1998-01-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-019-532-4
```

```
Query Match 89.4%; Score 2010.5; DB 4; Length 433;
Best Local Similarity 87.8%; Pred. No. 8.4e-155;
Matches 381; Conservative 30; Mismatches 22; Indels 1; Gaps 1;
```


QY 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 60
DB 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 60
QY 61 NANDPNGHGHVAGSVLNGTNGKGMAPQANLVFQSVMDNSGGLGLPSNVSTLPSQAYS 120
DB 61 NASDPNGHGHVAGSVLNGTNGKGMAPQANLVFQSVMDNSGGLGLPSNVSTLPSQAWN 119
QY 121 AGARIHTNSWGAPVNGAYTTDSRNVDDYVRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
DB 120 AGARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNGGTISAPGTAKNAI 179
QY 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAFGTIFLSARSSSLAPDSSF 240
DB 180 TVGATENYRPSFGSIADPNHIAQFSSRGATRDGRIKPDVTAEGTIFLSARSSSLAPDSSF 239
QY 241 WANHDSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLLKAALIAGATDGLGY 300
DB 240 WANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLLKAALIAGATDGLGY 299
QY 301 PSNGQGWGRVTLDKSLNVAFAVNETSSLTNOKATYSFTAQSGKPLKISLVMSDAPASTSA 360
DB 300 PSNGQGWGRVTLDKSLNVAFAVNETSSLTNOKATYSFTAQSGKPLKISLVMSDAPASTSA 359
QY 361 SYTLVNDLDELVITAPNGTKYVGNDFTPAYDNNWGRNNVENFINAPQSGTYTVEVOAYN 420
DB 360 SYTLVNDLDELVITAPNGTKYVGNDFTPAYDNNWGRNNVENFINAPQSGTYTVEVOAYN 419
QY 421 VPOGQPAFSLAIVN 434
DB 420 VPSGPORFSLAIVH 433

RESULT 9
US-09-338-746-4
; Sequence 4, Application US/09338746
; Patent No. 6638526
; GENERAL INFORMATION:
; APPLICANT: Deussen, Heinz-Josef
; APPLICANT: Olsen, Arne A.
; APPLICANT: Fatum, Tine M.
; APPLICANT: Roggen, Erwin L.
; TITLE OF INVENTION: A Polypeptide-Polymer Conjugate
; FILE REFERENCE: 5619.200-US
; CURRENT APPLICATION NUMBER: US/09/338,746
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: PA 1998 00809
; EARLIER FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: 60/091,461
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRI
; ORGANISM: Bacillus sp.
US-09-338-746-4

Query Match 89.4%; Score 2010.5; DB 4; Length 433;
Best Local Similarity 87.8%; Pred. No. 8.4e-155;
Matches 381; Conservative 30; Mismatches 22; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 60
DB 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 60
QY 61 NANDPNGHGHVAGSVLNGTNGKGMAPQANLVFQSVMDNSGGLGLPSNVSTLPSQAYS 120
DB 61 NASDPNGHGHVAGSVLNGTNGKGMAPQANLVFQSVMDNSGGLGLPSNVSTLPSQAWN 119
QY 121 AGARIHTNSWGAPVNGAYTTDSRNVDDYVRKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180

DB 120 AGARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNGGTISAPGTAKNAI 179
QY 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAFGTIFLSARSSSLAPDSSF 240
DB 180 TVGATENYRPSFGSIADPNHIAQFSSRGATRDGRIKPDVTAEGTIFLSARSSSLAPDSSF 239
QY 241 WANHDSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLLKAALIAGATDGLGY 300
DB 240 WANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLLKAALIAGATDGLGY 299
QY 301 PSNGQGWGRVTLDKSLNVAFAVNETSSLTNOKATYSFTAQSGKPLKISLVMSDAPASTSA 360
DB 300 PSNGQGWGRVTLDKSLNVAFAVNETSSLTNOKATYSFTAQSGKPLKISLVMSDAPASTSA 359
QY 361 SYTLVNDLDELVITAPNGTKYVGNDFTPAYDNNWGRNNVENFINAPQSGTYTVEVOAYN 420
DB 360 SYTLVNDLDELVITAPNGTKYVGNDFTPAYDNNWGRNNVENFINAPQSGTYTVEVOAYN 419
QY 421 VPOGQPAFSLAIVN 434
DB 420 VPSGPORFSLAIVH 433

RESULT 10
US-08-873-479-43
; Sequence 43, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Lynne, Christianson
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agtis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-873-479-43

Query Match 89.4%; Score 2010.5; DB 2; Length 635;
Best Local Similarity 87.8%; Pred. No. 1.4e-154;
Matches 381; Conservative 30; Mismatches 22; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 60
DB 203 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 262

Qy 61 NANDPNHGTHVAGSVLNGTSTKNGMAPOANLVFQSVMSDNGGLGSLPSNVSTLFSQAYS 120
Db 263 NASDPNGHGHVAGSVLNGAL-NKGMAPQANLVFQSVMSDNGGLGSLPSNVSTLFSQAWN 321
Qy 121 AGARIHNSWGAPVNGAYTTDSRVDDYVRKNDMAVLFAAGNEGNGGTISAPGTAKNAI 180
Db 322 AGARIHNSWGAPVNGAYTANSQVDEYVRNDMTVLFAAGNEGNGGTISAPGTAKNAI 381
Qy 181 TVGATENLRPSFGSYADNINHVAQFSRGTQDRI:KPDVMAPGTIFILSARSSLAPDSSF 240
Db 382 TVGATENLRPSFGSYADNINHVAQFSRGTQDRI:KPDVTAFTGTFILSARSSLAPDSSF 441
Qy 241 WANHDSKYVWGGTSMATPIVAGNVAQLREHF:KNGRITPKPSLLKAALIAGATDGLGY 300
Db 442 WANYSKIYVWGGTSMATPIVAGNVAQLREHF:KNGRITPKPSLLKAALIAGATDVGIGY 501
Qy 301 PSNGQWGRVTLDKSLNVAFVNETSSLSLNQKATYSTAQSGKPLKISLVWDAPASTSA 360
Db 502 PSQDQWGRVTLDKSLNVAVNEATALATQKATYSFQAQAGKPLKISLVWTDAPGSTTA 561
Qy 361 SVTLVNDLVLITAPNGTKYVGNDFTPYDNNWGDGRNNVFNINAPQSGTIVVEQAYN 420
Db 562 SYTLVNDLVLITAPNGTKYVGNDFSPYDNNWGDGRNNVFNINAPQSGTIVVEQAYN 621
Qy 421 VPOQPOAFSLAIVN 434
Db 622 VPSGPQRFSLAIVH 635

RESULT 11
US-09-512-251A-10

; Sequence 10, Application US/09512251A
; Patent No. 6555355

; GENERAL INFORMATION:

; APPLICANT: Hansen, Peter

; APPLICANT: Bauditz, Peter

; APPLICANT: Mikkelsen, Frank

; APPLICANT: Andersen, Kim

; TITLE OF INVENTION: Protease Variants and Compositions

; FILE REFERENCE: 5349.204-US

; CURRENT APPLICATION NUMBER: US/09/512,251A

; CURRENT FILING DATE: 2000-02-24

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 10

; LENGTH: 345

; TYPE: PRT

; ORGANISM: Bacillus

US-09-512-251A-10

Query Match 69.4%; Score 1562.5; DB 4; Length 345;

Best Local Similarity 92.5%; Pred. No. 1.1e-118;

Matches 294; Conservative 16; Mismatches 7; Indels 1; Gaps 1;

Qy 1 NDVARGIVKADVAQSSVGLYGGQGVAVADTGLDGTGRNDSMHEAFRGKITAIYALGRTN 60

Db 29 NDVARGIVKADVAQNFGLYGQGVAVADTGLDGTGRNDSMHEAFRGKITAIYALGRTN 88

Qy 61 NANDPNHGTHVAGSVLNGTSTKNGMAPOANLVFQSVMSDNGGLGSLPSNVSTLFSQAYS 120

Db 89 NANDPNHGTHVAGSVLNGAL-NKGMAPQANLVFQSVMSDNGGLGSLPSNVSTLFSQAYS 147

Qy 121 AGARIHNSWGAPVNGAYTTDSRVDDYVRKNDMAVLFAAGNEGNGGTISAPGTAKNAI 180

Db 148 AGARIHNSWGAPVNGAYTTDSRVDDYVRKNDMTLFAAGNEGPGSGTISAPGTAKNAI 207

Qy 181 TVGATENLRPSFGSYADNINHVAQFSRGTQDRI:KPDVMAPGTIFILSARSSLAPDSSF 240

Db 208 TVGATENLRPSFGSYADNINHVAQFSRGTQDRI:KPDVMAPGTIFILSARSSLAPDSSF 267

Qy 241 WANHDSKYVWGGTSMATPIVAGNVAQLREHF:KNGRITPKPSLLKAALIAGATDGLGY 300

Db 268 WANHDSKIYVWGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAGADVGLGF 327

Qy 301 PSNGQWGRVTLDKSLNV 318

Db 328 PNGNQWGRVTLDKSLNV 345

RESULT 12

US-09-515-150A-10

; Sequence 10, Application US/09515150A

; Patent No. 6558938

; GENERAL INFORMATION:

; APPLICANT: Hansen, Peter

; APPLICANT: Bauditz, Peter

; APPLICANT: Mikkelsen, Frank

; APPLICANT: Andersen, Kim

; TITLE OF INVENTION: Protease Variants and Compositions

; FILE REFERENCE: 5348.204-US

; CURRENT APPLICATION NUMBER: US/09/515,150A

; CURRENT FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 10

; LENGTH: 345

; TYPE: PRT

; ORGANISM: Bacillus

US-09-515-150A-10

Query Match 69.4%; Score 1562.5; DB 4; Length 345;

Best Local Similarity 92.5%; Pred. No. 1.1e-118;

Matches 294; Conservative 16; Mismatches 7; Indels 1; Gaps 1;

Qy 1 NDVARGIVKADVAQSSVGLYGGQGVAVADTGLDGTGRNDSMHEAFRGKITAIYALGRTN 60

Db 29 NDVARGIVKADVAQNFGLYGQGVAVADTGLDGTGRNDSMHEAFRGKITAIYALGRTN 88

Qy 61 NANDPNHGTHVAGSVLNGTSTKNGMAPOANLVFQSVMSDNGGLGSLPSNVSTLFSQAYS 120

Db 89 NANDPNHGTHVAGSVLNGAL-NKGMAPQANLVFQSVMSDNGGLGSLPSNVSTLFSQAYS 147

Qy 121 AGARIHNSWGAPVNGAYTTDSRVDDYVRKNDMAVLFAAGNEGPGGTISAPGTAKNAI 180

Db 148 AGARIHNSWGAPVNGAYTTDSRVDDYVRKNDMTLFAAGNEGPGGTISAPGTAKNAI 207

Qy 181 TVGATENLRPSFGSYADNINHVAQFSRGTQDRI:KPDVMAPGTIFILSARSSLAPDSSF 240

Db 208 TVGATENLRPSFGSYADNINHVAQFSRGTQDRI:KPDVMAPGTIFILSARSSLAPDSSF 267

Qy 241 WANHDSKYVWGGTSMATPIVAGNVAQLREHF:KNGRITPKPSLLKAALIAGATDGLGY 300

Db 268 WANHDSKIYVWGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAGADVGLGF 327

RESULT 13

US-09-196-281-13

; Sequence 13, Application US/09196281A

; Patent No. 6605458

; GENERAL INFORMATION:

; APPLICANT: Hansen, Peter K.

; APPLICANT: Bauditz, Peter

; APPLICANT: Mikkelsen, Frank

; TITLE OF INVENTION: Protease Variants And Compositions

; FILE REFERENCE: 5435.200-US

; CURRENT APPLICATION NUMBER: US/09/196,281A

; CURRENT FILING DATE: 1998-11-19

; EARLIER APPLICATION NUMBER: 1332/97

; EARLIER FILING DATE: 1997-11-21

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 13

```
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-09-196-281-13

Query Match      69.4%; Score 1562.5; DB 4; Length 345;
Best Local Similarity 92.5%; Pred. No. 1.1e-118;
Matches 294; Conservative 16; Mismatches 7; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 60
Db 29 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTN 88

QY 61 NADNENGHTHVAGSVGLNGTSKMGAPQANLVFQSVMDNSGGLGLPSNVSTLFSQAYS 120
Db 89 NADNENGHTHVAGSVGLNGAT-NKGAPQANLVFQSVMDNSGGLGLPANLQTLFSQAYS 147

QY 121 AGARIHTNSWGPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNNGGTISAPGTAKNAI 180
Db 148 AGARIHTNSWGPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 207

QY 181 TVGATENLRPSFGSYADNHNHVAQSSRGPTDGRIPKPDVMAPGTIFILSARSSLPDSSF 240
Db 208 TVGATENLRPSFGSYADNHNHVAQSSRGPTDGRIPKPDVMAPGTIFILSARSSLPDSSF 267

QY 241 WANHDSKYAYMGTSMATPIVAGNVAQLREHFINKRGITPKESLLKAAIAGATDIGLY 300
Db 268 WANHDSKYAYMGTSMATPIVAGNVAQLREHFVQNRGVTPEKESLLKAAIAGADVGLGF 327

QY 301 PSNGQGWGRVTLDKSLNV 318
Db 328 PSNGQGWGRVTLDKSLNV 345

RESULT 14
US-08-894-818B-1
; Sequence 1, Application US/08894818B
; Patent No. 6261822
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MITTA, Masanori
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,818B
; FILING DATE: 20-May-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03253
; FILING DATE: 07-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 323285/1995
; FILING DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
```

```
; REFERENCE/DOCKET NUMBER: TAKAKURA=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-818B-1

Query Match      20.0%; Score 450.5; DB 3; Length 659;
Best Local Similarity 31.4%; Pred. No. 3.1e-28;
Matches 133; Conservative 69; Mismatches 170; Indels 63; Gaps 15;

QY 8 VKADVAQSSYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITAIY-ALGRTNNDNDPN 66
Db 145 IGADTVMSNLGYDGSVVAIVDTGIDAN-----HPDLKKGKIVGYDAVNGRSTPYDDQ 198

QY 67 GHGTHVAGSVLNGTSNK---GMAPQANLVFQSVM---DSNGLGLGLPSNVSTLFSQAYS 121
Db 199 GHGTHVAGSVLNGTSNK---GMAPQANLVFQSVM---DSNGLGLGLPSNVSTLFSQAYS 258

QY 122 GARIHTNSWGPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNNGGTISAPGTAKNAI 181
Db 259 GIRVINLSLSSQSSDSDTSLSQAVNNAMDAGIVVCVAGNSGPNVTYVGSPPAAASKVIT 318

QY 182 VGATENLRPSFGSYADNHNHVAQSSRGPTDGRIPKPDVMAPGTIFILSARSSLPDSSF 241
Db 319 VGA-----VDSNDNIASFSSRGPTADGRIPKPDVMAPGTIFILSARSSLPDSSF 364

QY 242 ANHDSKYAYMGTSMATPIVAGNVAQLREHFINKRGITP---KPSLLKAAIAGATDIG- 297
Db 365 TPINDYTKASGTSMATPHVSGVGAALIQ---AHPSTPDKVKTALETADIVAPKEIAD 421

QY 298 LGYPSNGQGWGRVTLDKSL---NVAFVNETSLSSTNQKATYSTAQSGLKSLIVWSDA 354
Db 422 IAY-----GAGRVNVYKAIKYDDYAKLFTGVSADKGSATHTTDPVSGATFVTATLYWD-- 474

QY 355 PASTSASVTLVNDLILVITAPNGTKYVGNDFTPAYDNNMDGRNNDVNFVINAPOSQGYTV 414
Db 475 -----TGSSDIDLVDPNPNE-VDYSYTAIY-----GFEKVGYYNPTAGTWTV 517

QY 415 EVQAYNVPPQGFAPFLAIVN 434
Db 518 KWSY---KGAANYQVDWS 534

RESULT 15
US-09-445-472-12
; Sequence 12, Application US/09445472
; Patent No. 6358726
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/09/445,472
; CURRENT FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Thermococcus celer
US-09-445-472-12
```

Query Match 20.0%; Score 450.5; DB 4; Length 659;
Best Local Similarity 31.4%; Pred. No. 3.1e-28;
Matches 138; Conservative 69; Mismatches 170; Indels 63; Gaps 15;

| | | | |
|----|-----|---|-----|
| Qy | 8 | VKADVAQSSVGLYGQGVAVADTGLDTGRNDSSHEAFRGKITAIY-ALGRINNANDPN | 66 |
| Db | 145 | IGADTVNSLGYDGGVVAIVDTGTIDAN-----HPDLKGVIGYDAVNGRSTPYDDQ | 198 |
| Qy | 67 | GHGTHVAGSVLNGTSNK---GMAPOANLVFOSVM--DSNGGLGLPLSNVSTLFSQAYSA | 121 |
| Db | 199 | GHGTHVAGIVAGTGSVNSQYIGVAPGAKLVGVKVLGADGSGSVSTIIAGVDVVQNKDKY | 258 |
| Qy | 122 | GARIHTNSWGAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNCGTISAPGTAKNAIT | 181 |
| Db | 259 | GIRVINLSLSSOSSDGTSLSCAVANNANDAGIVVCVRAAGNSGPNTYTVGSPAASKVIT | 318 |
| Qy | 182 | VGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAPGTFFILSARSLAPDSSFW | 241 |
| Db | 319 | VGA-----VDSNDNIASFSSRGPTADGRLKPEVVAPGVVDIIAPRAS---GTSMG | 364 |
| Qy | 242 | ANHDSKYAYMGTSMATPIVAGNVAQLREHFINKRGITP---KPSLLKAALIAGATDIG- | 297 |
| Db | 365 | TPINDYTKASGTSMATPHVSGVGCALILQ---AHPSWTDKVKTALITADIAPKEIAD | 421 |
| Qy | 298 | LGYPSGNQGWGRVTLDKSL---NVAFYNETSSLSTNQKATYSFTAQSGKPLKISLVWSDA | 354 |
| Db | 422 | IAY-----GAGRVNVYKAIKYDDYAKLITFTGSVADKGSATHTFDVSGATFVTATLYWD-- | 474 |
| Qy | 355 | PASTSASVTLVNDLILVITAPNGTKYVGNDFAPYDNNWDCGRNNVENFINAPQSGTYTV | 414 |
| Db | 475 | -----TGSSDIDLXYDPNGNE-VDYSYTAY-----GFEKVGYYNPTAGTWTIV | 517 |
| Qy | 415 | EVOAYNVPGQPQAFSLAIVN | 434 |
| Db | 518 | KVUSY---XGAANYQVDVVS | 534 |

Search completed: March 31, 2004, 16:11:56
Job time : 15.0215 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 16:09:09 ; Search time 33.2177 Seconds
(without alignments)
3418.697 Million cell updates/sec

Title: US-09-985-689A-6
Perfect score: 2230
Sequence: 1 NDVARGIVKADVAQSSGLY.....EVQYVNPQGFQAFSLAVN 434

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 261661801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/1/pubpaa/US09D_NEW_PUB.pep.*
13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/1/pubpaa/US10D_NEW_PUB.pep.*
17: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|---------------------|
| 1 | 2250 | 100.0 | 434 | 10 | US-09-985-689A-6 |
| 2 | 2143 | 95.2 | 434 | 10 | US-09-985-689A-1 |
| 3 | 2143 | 95.2 | 434 | 15 | US-10-385-662-2 |
| 4 | 2130 | 94.7 | 434 | 10 | US-09-985-689A-2 |
| 5 | 2116.5 | 94.1 | 433 | 10 | US-09-985-689A-7 |
| 6 | 2024.5 | 90.0 | 433 | 10 | US-09-985-689A-5 |
| 7 | 2020.5 | 89.8 | 433 | 10 | US-09-985-689A-3 |
| 8 | 2005.5 | 89.1 | 433 | 10 | US-09-985-689A-4 |
| 9 | 1562.5 | 69.4 | 345 | 14 | US-10-336-324-10 |
| 10 | 1562.5 | 69.4 | 345 | 14 | US-10-403-105-13 |
| 11 | 450.5 | 20.0 | 659 | 13 | US-10-090-624-12 |
| 12 | 404 | 18.0 | 412 | 13 | US-10-090-624-1 |
| 13 | 404 | 18.0 | 522 | 13 | US-10-090-624-4 |
| 14 | 404 | 18.0 | 654 | 13 | US-10-090-624-16 |
| 15 | 362 | 16.1 | 1208 | 14 | US-10-156-761-13251 |

| | | | | | | |
|----|-------|------|-------|----|---------------------|-------------------|
| 16 | 357 | 15.9 | 1079 | 14 | US-10-112-488-39 | Sequence 39, Appl |
| 17 | 349.5 | 15.5 | 1139 | 14 | US-10-156-761-10856 | Sequence 10856, A |
| 18 | 339 | 15.1 | 1237 | 14 | US-10-314-657-4 | Sequence 4, Appl |
| 19 | 316.5 | 14.1 | 1398 | 13 | US-10-090-624-6 | Sequence 6, Appl |
| 20 | 297.5 | 13.2 | 519 | 15 | US-10-084-846A-114 | Sequence 114, App |
| 21 | 297.5 | 13.2 | 19725 | 15 | US-10-084-846A-4 | Sequence 4, Appl |
| 22 | 282.5 | 12.6 | 595 | 10 | US-09-927-827-59 | Sequence 59, Appl |
| 23 | 274 | 12.2 | 580 | 10 | US-09-966-921A-2 | Sequence 55, Appl |
| 24 | 272 | 12.1 | 418 | 9 | US-09-966-921A-2 | Sequence 2, Appl |
| 25 | 266 | 11.8 | 1101 | 14 | US-10-156-761-12934 | Sequence 12934, A |
| 26 | 264 | 11.7 | 271 | 10 | US-09-813-408-2 | Sequence 2, Appl |
| 27 | 264 | 11.7 | 280 | 14 | US-10-209-812-2 | Sequence 2, Appl |
| 28 | 260.5 | 11.6 | 397 | 10 | US-09-773-334A-5 | Sequence 5, Appl |
| 29 | 258 | 11.5 | 368 | 12 | US-10-344-231-3 | Sequence 3, Appl |
| 30 | 258 | 11.5 | 271 | 14 | US-10-242-549-54 | Sequence 54, Appl |
| 31 | 258 | 11.5 | 271 | 14 | US-10-242-549-56 | Sequence 56, Appl |
| 32 | 258 | 11.5 | 271 | 14 | US-10-242-549-60 | Sequence 60, Appl |
| 33 | 258 | 11.5 | 275 | 14 | US-10-324-152-11 | Sequence 11, Appl |
| 34 | 257 | 11.4 | 269 | 8 | US-08-322-678-10 | Sequence 10, Appl |
| 35 | 257 | 11.4 | 269 | 9 | US-09-837-235-16 | Sequence 16, Appl |
| 36 | 257 | 11.4 | 269 | 9 | US-09-060-854B-6 | Sequence 6, Appl |
| 37 | 257 | 11.4 | 269 | 9 | US-09-975-139-1 | Sequence 1, Appl |
| 38 | 257 | 11.4 | 269 | 9 | US-09-976-414-8 | Sequence 8, Appl |
| 39 | 257 | 11.4 | 269 | 10 | US-09-736-116-49 | Sequence 49, Appl |
| 40 | 257 | 11.4 | 269 | 13 | US-10-075-907-1 | Sequence 1, Appl |
| 41 | 257 | 11.4 | 269 | 13 | US-10-075-895-1 | Sequence 1, Appl |
| 42 | 257 | 11.4 | 289 | 14 | US-10-033-325-6 | Sequence 6, Appl |
| 43 | 257 | 11.4 | 289 | 14 | US-10-209-812-3 | Sequence 3, Appl |
| 44 | 257 | 11.4 | 269 | 14 | US-10-104-693-4 | Sequence 4, Appl |
| 45 | 257 | 11.4 | 269 | 14 | US-10-228-572-6 | Sequence 6, Appl |

ALIGNMENTS

RESULT 1

US-09-985-689A-6
; Sequence 6, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985,689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-6

Query Match 100.0%; Score 2250; DB 10; Length 434;
Best Local Similarity 100.0%; Pred. No. 3e-189;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQSSGLYGGQGVAVADTGLDGRNDSMHEAFRGKTIIVALGRTN 60

DB 1 NDVARGIVKADVAQSSGLYGGQGVAVADTGLDGRNDSMHEAFRGKTIIVALGRTN 60

QY 61 NANDPNHGHTHVAGSVLNGTSTNKGMPQANLVFQVMSDNGGLGGLPSNVTLSQAYS 120

Db 61 NNDPNGHGHVAGSVLNGTGNKGMAPQANLVFQSVMDNSGGLGGLFNSVTLFQAYS 120
QY 121 AGARHTNSWGAAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNCGTISAPGTAKNAI 180
Db 121 AGARHTNSWGAAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNCGTISAPGTAKNAI 180
QY 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAAGTFFILSARSSLPDSSF 240
Db 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAAGTFFILSARSSLPDSSF 240
QY 241 WANHDSKYAYMGTSMATPIVAGNVAQLREHFIKRGITPKPSLLKAALIAGATDGLGY 300
Db 241 WANHDSKYAYMGTSMATPIVAGNVAQLREHFIKRGITPKPSLLKAALIAGATDGLGY 300
QY 301 PSNGQGWGRVTLDKSLNVAFVNETSLTNQKATYSFTAQSGKPLKISLVWSDAPASTSA 360
Db 301 PSNGQGWGRVTLDKSLNVAFVNETSLTNQKATYSFTAQSGKPLKISLVWSDAPASTSA 360
QY 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGDGRNNVNFINAPOSGETTYVEQAYN 420
Db 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGDGRNNVNFINAPOSGETTYVEQAYN 420
QY 421 VPQGPQAFSLAIYN 434
Db 421 VPQGPQAFSLAIYN 434

RESULT 2

US-09-985-689A-1

; Sequence 1, Application US/09985689A

; Publication No. US20030022351A1

; GENERAL INFORMATION:

; APPLICANT: HATADA, YUJI

; APPLICANT: OGAWA, AKINORI

; APPLICANT: KAGETAMA, YASUSHI

; APPLICANT: SATO, TSUYOSHI

; APPLICANT: ARAKI, HIROYUKI

; APPLICANT: SUMITOMO, NOBUYUKI

; APPLICANT: OKUDA, MITSUYOSHI

; APPLICANT: SAEKI, KATSUHISA

; TITLE OF INVENTION: Alkaline proteases

; FILE REFERENCE: 215483US0

; CURRENT APPLICATION NUMBER: US/09/985,689A

; PRIOR FILING DATE: 2002-07-01

; PRIOR APPLICATION NUMBER: JP P2000-355166

; PRIOR FILING DATE: 2000-11-22

; PRIOR APPLICATION NUMBER: JP P2001-114048

; PRIOR FILING DATE: 2001-04-12

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 1

; LENGTH: 434

; TYPE: PRT

; ORGANISM: Bacillus sp.

US-09-985-689A-1

Query Match 95.2%; Score 2143; DB 10; Length 434;
Best Local Similarity 93.5%; Pred. No. 7.9e-180;
Matches 406; Conservative 19; Mismatches 9; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAOSSYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITAIYALGRTN 60
Db 1 NDVARGIVKADVAOSSYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITAIYALGRTN 60
QY 61 NNDPNGHGHVAGSVLNGTGNKGMAPQANLVFQSVMDNSGGLGGLFNSVTLFQAYS 120
Db 61 NNDPNGHGHVAGSVLNGTGNKGMAPQANLVFQSVMDNSGGLGGLFNSVTLFQAYS 120
QY 121 AGARHTNSWGAAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNCGTISAPGTAKNAI 180
Db 121 AGARHTNSWGAAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNCGTISAPGTAKNAI 180

QY 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAAGTFFILSARSSLPDSSF 240
Db 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAAGTFFILSARSSLPDSSF 240
QY 241 WANHDSKYAYMGTSMATPIVAGNVAQLREHFIKRGITPKPSLLKAALIAGATDGLGY 300
Db 241 WANHDSKYAYMGTSMATPIVAGNVAQLREHFIKRGITPKPSLLKAALIAGATDGLGY 300
QY 301 PSNGQGWGRVTLDKSLNVAFVNETSLTNQKATYSFTAQSGKPLKISLVWSDAPASTSA 360
Db 301 PSNGQGWGRVTLDKSLNVAFVNETSLTNQKATYSFTAQSGKPLKISLVWSDAPASTSA 360
QY 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGDGRNNVNFINAPOSGETTYVEQAYN 420
Db 361 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGDGRNNVNFINAPOSGETTYVEQAYN 420
QY 421 VPQGPQAFSLAIYN 434
Db 421 VPQGPQAFSLAIYN 434

RESULT 3
US-10-385-662-2
; Sequence 2, Application US/10385662
; Publication No. US2004002432A1
; GENERAL INFORMATION:
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: SAITO, KAZUHIRO
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: IZAWA, YOSHIFUMI
; APPLICANT: SAEKI, KATSUHISA
; APPLICANT: KOBAYASHI, TOHRU
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: Alkaline protease
; FILE REFERENCE: 234938US0
; CURRENT APPLICATION NUMBER: US/10/385,662
; CURRENT FILING DATE: 2003-03-12
; PRIOR FILING DATE: 2002-08-1428
; PRIOR APPLICATION NUMBER: JP 2002-165987
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: JP 2002-304230
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: JP 2002-304231
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp. KSM-KP43
US-10-385-662-2

Query Match 95.2%; Score 2143; DB 15; Length 434;
Best Local Similarity 93.5%; Pred. No. 7.9e-180;
Matches 406; Conservative 19; Mismatches 9; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAOSSYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITAIYALGRTN 60
Db 1 NDVARGIVKADVAOSSYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITAIYALGRTN 60
QY 61 NNDPNGHGHVAGSVLNGTGNKGMAPQANLVFQSVMDNSGGLGGLFNSVTLFQAYS 120
Db 61 NNDPNGHGHVAGSVLNGTGNKGMAPQANLVFQSVMDNSGGLGGLFNSVTLFQAYS 120
QY 121 AGARHTNSWGAAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNCGTISAPGTAKNAI 180
Db 121 AGARHTNSWGAAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNCGTISAPGTAKNAI 180
QY 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAAGTFFILSARSSLPDSSF 240
Db 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAAGTFFILSARSSLPDSSF 240

QY 241 WANHDSKYAYMGSTSMATPIVAGNVAQLREHFIRKNGITPKPSLLKAALIAGATDGLGY 300
Db 241 WANHDSKYAYMGSTSMATPIVAGNVAQLREHFIRKNGITPKPSLLKAALIAGATDGLGY 300
QY 301 PSNGQGWGRVTLDKSLNVAQVNETSSLTNQKATYSFTAQSGKPLKISLVWSDAPASTSA 360
Db 301 PSNGQGWGRVTLDKSLNVAQVNETSSLTNQKATYSFTAQSGKPLKISLVWSDAPASTSA 360
QY 361 SVTLVNDLDLVTAPNGTKYVGNDFTPADYNNWGRNNVNFVFINAPQSGTYTVEVOAYN 420
Db 361 SVTLVNDLDLVTAPNGTKYVGNDFTPADYNNWGRNNVNFVFINAPQSGTYTVEVOAYN 420
QY 421 VPQGPQAFSLAIYN 434
Db 421 VPVGPQTFLAIYN 434

RESULT 4
US-09-985-689A-2
; Sequence 2, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985, 689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-2

Query Match 94.7%; Score 2130; DB 10; Length 434;
Best Local Similarity 92.4%; Pred. No. 1.1e-178;
Matches 401; Conservative 24; Mismatches 9; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTCGRNDSMHEAFRGKITAIYALGRTN 60
Db 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTCGRNDSMHEAFRGKITAIYALGRTN 60
QY 61 NNDPNHGHTHVAGSVLNGTSGKGAPOANLVFQSDMSNGGLGGLPSNVSTLFSQAYS 120
Db 61 NNDPNHGHTHVAGSVLNGTSGKGAPOANLVFQSDMSNGGLGGLPSNLQTLFSQAYS 120
QY 121 AGARIHTNSGAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNNGGTISAPGTAKNAI 180
Db 121 AGARIHTNSGAAVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPNNGGTISAPGTAKNAI 180
QY 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAFGTIFLSARSLAPDSFP 240
Db 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSFP 240
QY 241 WANHDSKYAYMGSTSMATPIVAGNVAQLREHFIRKNGITPKPSLLKAALIAGATDGLGY 300
Db 241 WANHDSKYAYMGSTSMATPIVAGNVAQLREHFIRKNGITPKPSLLKAALIAGADVGLGY 300
QY 301 PSNGQGWGRVTLDKSLNVAQVNETSSLTNQKATYSFTAQSGKPLKISLVWSDAPASTSA 360

Db 301 PSNGQGWGRVTLDKSLNVAQVNETSSLTNQKATYSFTAQSGKPLKISLVWSDAPASTTA 360
QY 361 SVTLVNDLDLVTAPNGTKYVGNDFTPADYNNWGRNNVNFVFINAPQSGTYTVEVOAYN 420
Db 361 SVTLVNDLDLVTAPNGTKYVGNDFTPADYNNWGRNNVNFVFINAPQSGTYTVEVOAYN 420
QY 421 VPQGPQAFSLAIYN 434
Db 421 VPVGPQAFSLAIYN 434

RESULT 5
US-09-985-689A-7
; Sequence 7, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985, 689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 7
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-7

Query Match 94.1%; Score 2116.5; DB 10; Length 433;
Best Local Similarity 92.2%; Pred. No. 1.7e-177;
Matches 400; Conservative 21; Mismatches 12; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTCGRNDSMHEAFRGKITAIYALGRTN 60
Db 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTCGRNDSMHEAFRGKITAIYALGRTN 60
QY 61 NNDPNHGHTHVAGSVLNGTSGKGAPOANLVFQSDMSNGGLGGLPSNVSTLFSQAYS 120
Db 61 NNDPNHGHTHVAGSVLNGTSGKGAPOANLVFQSDMSNGGLGGLPANLQTLFSQAYS 119
QY 121 AGARIHTNSGAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNNGGTISAPGTAKNAI 180
Db 120 AGARIHTNSGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPNNGGTISAPGTAKNAI 179
QY 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAFGTIFLSARSLAPDSFP 240
Db 180 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSFP 239
QY 241 WANHDSKYAYMGSTSMATPIVAGNVAQLREHFIRKNGITPKPSLLKAALIAGATDGLGY 300
Db 240 WANHDSKYAYMGSTSMATPIVAGNVAQLREHFIRKNGITPKPSLLKAALIAGADVGLGF 299
QY 301 PSNGQGWGRVTLDKSLNVAQVNETSSLTNQKATYSFTAQSGKPLKISLVWSDAPASTSA 360
Db 300 PSNGQGWGRVTLDKSLNVAQVNETSSLTNQKATYSFTAQSGKPLKISLVWSDAPGSTTA 359
QY 361 SVTLVNDLDLVTAPNGTKYVGNDFTPADYNNWGRNNVNFVFINAPQSGTYTVEVOAYN 420
Db 360 SVTLVNDLDLVTAPNGTKYVGNDFTPADYNNWGRNNVNFVFINAPQSGTYTVEVOAYN 419
QY 421 VPQGPQAFSLAIYN 434

Db 420 VPVSPOTFSLAIVH 433

RESULT 6

US-09-985-689A-5

; Sequence 5, Application US/09985689A

; Publication No. US20030022351A1

; GENERAL INFORMATION:

; APPLICANT: HATADA, YUJI

; APPLICANT: OGAWA, AKINORI

; APPLICANT: KAGEYAMA, YASUSHI

; APPLICANT: SATO, TSUYOSHI

; APPLICANT: ARAKI, HIROYUKI

; APPLICANT: SUMITOMO, NOBUYUKI

; APPLICANT: OKUDA, MITSUYOSHI

; APPLICANT: SAEKI, KATSUHISA

; TITLE OF INVENTION: Alkaline proteases

; FILE REFERENCE: 215483US0

; CURRENT APPLICATION NUMBER: US/09/985,689A

; PRIOR FILING DATE: 2002-07-01

; PRIOR APPLICATION NUMBER: JP P2000-355166

; PRIOR FILING DATE: 2000-11-22

; PRIOR APPLICATION NUMBER: JP P2001-114048

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 5

; LENGTH: 433

; TYPE: PRT

; ORGANISM: Bacillus sp.

US-09-985-689A-5

Query Match

Best Local Similarity 90.0%; Score 2024.5; DB 10; Length 433;

Matches 384; Conservative 28; Mismatches 21; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYQGVVAVADTGLDGTGRNDSMHEAFRGKITAIYALGRIN 60

Db 1 NDVARGIVKADVAQNNYGLYQGVVAVADTGLDGTGRNDSMHEAFRGKITAIYALGRIN 60

QY 61 NANDPNHGHTHVAGSVLNGTSGKGMAPQANLVFQSVMDNSGGLGGLPSNVSTLFSQAYS 120

Db 61 NANDPNHGHTHVAGSVLGNAL-NKGMAPQANLVFQSVMDNSGGLGGLPSNVSTLFSQAWN 119

QY 121 AGARIHNSWGAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNNGGTISAPGTAKNAI 180

Db 120 AGARIHNSWGAPVNGAYTANSRQVDEYVRNDMTVLFAAGNEGPNNGGTISAPGTAKNAI 179

QY 181 TVGATENLRPSFGSYADNINHVAFQSSRGPTKGRIPKDVMAFGTIFLSARSLAPDSF 240

Db 180 TVGATENYRPSFGSLADNPNHIAQFSSRGATRDGRIPKDVTAFTPTILSARSLAPDSF 239

QY 241 WANHDSKYVMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLLKAALIAGATDGLGY 300

Db 240 WANYNSKYVMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLLKAALIAGATDGLGY 299

QY 301 PSNGQGRVRLDKSLNVAFVNETSSLTNQKATYSTAQSGLPKISLVWSDAPASTSA 360

Db 300 PSNGQGRVRLDKSLNVAFVNEATALATGQKATYSTAQSGLPKISLVWSDAPASTSA 359

QY 361 SVTLVNDLDLIVITAPNGTKVYVNDFTAPYDNNWGRNVENVFINAPQSGTIVTVEQAYN 420

Db 360 SYTLVNDLDLIVITAPNGTKVYVNDFTAPYDNNWGRNVENVFINAPQSGTIVTVEQAYN 419

QY 421 VPQGPQAFSLAIVH 434

Db 420 VPQGPQAFSLAIVH 433

RESULT 7

US-09-985-689A-3

; Sequence 3, Application US/09985689A

; Publication No. US20030022351A1

; GENERAL INFORMATION:

; APPLICANT: HATADA, YUJI

; APPLICANT: OGAWA, AKINORI

; APPLICANT: KAGEYAMA, YASUSHI

; APPLICANT: SATO, TSUYOSHI

; APPLICANT: ARAKI, HIROYUKI

; APPLICANT: SUMITOMO, NOBUYUKI

; APPLICANT: OKUDA, MITSUYOSHI

; APPLICANT: SAEKI, KATSUHISA

; TITLE OF INVENTION: Alkaline proteases

; FILE REFERENCE: 215483US0

; CURRENT APPLICATION NUMBER: US/09/985,689A

; PRIOR FILING DATE: 2002-07-01

; PRIOR APPLICATION NUMBER: JP P2000-355166

; PRIOR FILING DATE: 2000-11-22

; PRIOR APPLICATION NUMBER: JP P2001-114048

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 3

; LENGTH: 433

; TYPE: PRT

; ORGANISM: Bacillus sp.

US-09-985-689A-3

Query Match

Best Local Similarity 89.8%; Score 2020.5; DB 10; Length 433;

Matches 383; Conservative 28; Mismatches 22; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYQGVVAVADTGLDGTGRNDSMHEAFRGKITAIYALGRIN 60

Db 1 NDVARGIVKADVAQNNYGLYQGVVAVADTGLDGTGRNDSMHEAFRGKITAIYALGRIN 60

QY 61 NANDPNHGHTHVAGSVLNGTSGKGMAPQANLVFQSVMDNSGGLGGLPSNVSTLFSQAYS 120

Db 61 NANDPNHGHTHVAGSVLGNAL-NKGMAPQANLVFQSVMDNSGGLGGLPSNVSTLFSQAWN 119

QY 121 AGARIHNSWGAPVNGAYTTDSRNVDYVRKNDMAVLFAAGNEGPNNGGTISAPGTAKNAI 180

Db 120 AGARIHNSWGAPVNGAYTANSRQVDEYVRNDMTVLFAAGNEGPNNGGTISAPGTAKNAI 179

QY 181 TVGATENLRPSFGSYADNINHVAFQSSRGPTKGRIPKDVMAFGTIFLSARSLAPDSF 240

Db 180 TVGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIPKDVTAFTPTILSARSLAPDSF 239

QY 241 WANHDSKYVMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLLKAALIAGATDGLGY 300

Db 240 WANYNSKYVMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLLKAALIAGATDGLGY 299

QY 301 PSNGQGRVRLDKSLNVAFVNETSSLTNQKATYSTAQSGLPKISLVWSDAPASTSA 360

Db 300 PSNGQGRVRLDKSLNVAFVNEATALATGQKATYSTAQSGLPKISLVWSDAPASTSA 359

QY 361 SVTLVNDLDLIVITAPNGTKVYVNDFTAPYDNNWGRNVENVFINAPQSGTIVTVEQAYN 420

Db 360 SYTLVNDLDLIVITAPNGTKVYVNDFTAPYDNNWGRNVENVFINAPQSGTIVTVEQAYN 419

QY 421 VPQGPQAFSLAIVH 434

Db 420 VPQGPQAFSLAIVH 433

RESULT 8

US-09-985-689A-4

; Sequence 4, Application US/09985689A

; Publication No. US20030022351A1

; GENERAL INFORMATION:

; APPLICANT: HATADA, YUJI

; APPLICANT: OGAWA, AKINORI

; APPLICANT: KAGEYAMA, YASUSHI

; APPLICANT: SATO, TSUYOSHI

; APPLICANT: ARAKI, HIROYUKI


```

; SEQ ID NO 10
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-10-336-324-10

Query Match          69.4%; Score 1562.5; DB 14; Length 345;
Best Local Similarity 92.5%; Pred. No. 6.6e-129;
Matches 294; Conservative 16; Mismatches 7; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGGOQVAVADTGLDTCGRNDSSMHEAFRGKITAYALGRTN 60
DB 1 .....:.....:.....:.....:.....:.....:.....:.....:.....:.....:
29 NDVARGIVKADVAQNNGFLYGGQIVAVADTGLDTCGRNDSSMHEAFRGKITAYALGRTN 88
QY 61 NANDPNHGTHVAGSVLNGTNSKMGAPQANLVFQSVMDNSNGGLGGLPSNVSTLFSQAYS 120
DB 1 .....:.....:.....:.....:.....:.....:.....:.....:.....:
89 NANDPNHGTHVAGSVLGNAT-NKMGAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 147
QY 121 AGARIHTNSWGPVNGAYTTDSRNVDDYVRKNDMAVLFAAGNEGPGNGGTTISAPGTAKNAI 180
DB 148 AGARIHTNSWGPVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPGSGTILISAPGTAKNAI 207
QY 181 TVGATENLRPSFGSYADNINHVAQFSRGFTKQGRIKPDVMAFGTTFILSARSSLAPDSF 240
DB 208 TVGATENLRPSFGSYADNINHVAQFSRGFTRDGRICKPDVMAFGTTFILSARSSLAPDSF 267
QY 241 WANHDSKIAYMGGTSMATPIVAGNVAQLREHFHKRGITPKPSLLKAALITAGATDGLGY 300
DB 268 WANHDSKIAYMGGTSMATPIVAGNVAQLREHFVKRNGVTPKPSLLKAALITAGAADVGLGF 327
QY 301 PSGNQCGWRVTLDKSLNV 318
DB 328 PNGNQCGRVTLTDKSLNV 345

RESULT 10
US-10-403-105-13
; Sequence 13, Application US/10403105
; Publication No. US20030180933A1
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter K.
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; TITLE OF INVENTION: Protease Variants And Compositions
; FILE REFERENCE: 5435.200-US
; CURRENT APPLICATION NUMBER: US/10/403,105
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US/09/196,281A
; PRIOR FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 1332/97
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-10-403-105-13

Query Match          69.4%; Score 1562.5; DB 14; Length 345;
Best Local Similarity 92.5%; Pred. No. 6.6e-129;
Matches 294; Conservative 16; Mismatches 7; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGGOQVAVADTGLDTCGRNDSSMHEAFRGKITAYALGRTN 60
DB 29 NDVARGIVKADVAQNNGFLYGGQIVAVADTGLDTCGRNDSSMHEAFRGKITAYALGRTN 88
QY 61 NANDPNHGTHVAGSVLNGTNSKMGAPQANLVFQSVMDNSNGGLGGLPSNVSTLFSQAYS 120
DB 89 NANDPNHGTHVAGSVLGNAT-NKMGAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 147
QY 121 AGARIHTNSWGPVNGAYTTDSRNVDDYVRKNDMAVLFAAGNEGPGNGGTTISAPGTAKNAI 180

```

Db 148 AGAIIHTNSGAPVNGAYTTDSNRVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 207
Qy 181 TVGATENLRPSFGSYADININHVAFSSRGPTKGRIPKDPVMAPTGTFILSARSSLAPDSF 240
Db 208 TVGATENLRPSFGSYADININHVAFSSRGPTKGRIPKDPVMAPTGTFILSARSSLAPDSF 267
Qy 241 WANHDSKYAYMGTSMTPTVAGNVAQLREHFKNGITPKPSLLKAALIAGATDGLCY 300
Db 268 WANHDSKYAYMGTSMTPTVAGNVAQLREHFKNGITPKPSLLKAALIAGATDGLCY 327
Qy 301 PSNGQWGRVTLDKSLNV 318
Db 328 PSNGQWGRVTLDKSLNV 345

RESULT 11
US-10-090-624-12
; Sequence 12, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Thermococcus celer
US-10-090-624-12

Query Match 20.0%; Score 450.5; DB 13; Length 659;
Best Local Similarity 31.4%; Pred. No. 1.2e-30;
Matches 138; Conservative 69; Mismatches 170; Indels 63; Gaps 15;

Qy 8 VKADVAOSSGLYCGQGVAVADTGLDTRNDSSMHEAFRGKITAIY-ALGRTNNDN 66
Db 145 IGADTVWNSLGYDGSVVVAIVDTGIDAN-----HPDLKGVIGWYDAVNGRSTPYDQ 198
Qy 67 GHGTHVAGSVLNGTGNK---GMAPOANLVFQSVN---DSNGLGLGPSNVSTLFSQAYSA 121
Db 199 GHGTHVAGSVLNGTGNK---GMAPOANLVFQSVN---DSNGLGLGPSNVSTLFSQAYSA 121
Qy 122 GARHTNSGAPVNGAYTTDSNRVDYVRKNDMAVLPAAGNEGPGSGTISAPGTAKNAIT 181
Db 259 GIRVINLSLSSQSSDGTDSLQAVNNAWDAGIIVCVVAGNSGPNVTYVTSPPAAASKVIT 318
Qy 182 VGATENLRPSFGSYADININHVAFSSRGPTKGRIPKDPVMAPTGTFILSARSSLAPDSF 241
Db 319 VGA-----VDSDNIAFSSRGPTADGRKPEVAPGVADIIAPRAS---GTSMG 364
Qy 242 ANHDSKYAYMGTSMTPTVAGNVAQLREHFKNGITP---KPSLLKAALIAGATDGLG- 297
Db 365 TPINDYYTKASGTSMTAPVHSGVGLITLQ---AHPSTWTPDKVTALITADIIVAPKEIAD 421
Qy 298 LGVPSGNQWGRVTLDKSL---NVAFVNETSSLTNQKATYSFTAQSGKPLKISLWSDA 354
Db 422 IAY-----GAGRVNVAIKYDDYAKLTFTGSDVADKGSATHTPDVSGATFVTATLYWD- 474
Qy 355 PASTSASVTLNLDLVTAPNGTKYVGNDFTAIPYDNNWDRNNVNFVFINAPQSGTYTV 414
Db 475 -----TGSSDIDLVLDPNGNE-VDSYATAY-----GFEKVGYNPTAGTIVT 517

Qy 415 EVQAYNVPPQGPQAPSLAIVN 434
Db 518 KWSY---KGAANYQVDVWS 534

RESULT 12
US-10-090-624-1
; Sequence 1, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-10-090-624-1

Query Match 18.0%; Score 404; DB 13; Length 412;
Best Local Similarity 30.0%; Pred. No. 7.2e-27;
Matches 131; Conservative 63; Mismatches 168; Indels 74; Gaps 16;

Qy 18 GLYCGQGVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTNAN-----DPNGHGTH 71
Db 22 GYDGSGITIGITGID-----ASHPDLOGKV-----IGWDFVNGRSYPYDDHGHGTH 70
Qy 72 VAGSVLNGTGN---KMAPQANLVFQSVN---DSNGLGLGPSNVSTLFSQAYSAGARI 125
Db 71 VASIAAGTGAASNGKYKGMAPGAKLAGIKVLGADGSGSISTIIKGVEMAVDNKDKYGIK 130
Qy 126 HTNSGAPVNGAYTTDSNRVDYVRKNDMAVLPAAGNEGPGSGTISAPGTAKNAITVGA 185
Db 131 INLSLSSQSSDGTDSLQAVNNAWDAGIIVCVVAGNSGPNVTYVTSPPAAASKVITVGA- 189
Qy 186 ENLRPSFGSYADININHVAFSSRGPTKGRIPKDPVMAPTGTFILSARSSLAPDSF 245
Db 190 -----VDKYDVITSFSSRGPTADGRKPEVAPGNWIIAARAS---GTSMGQPIN 236
Qy 246 SKIAYMGTSMTPTVAGNVAQLREHFKNGITP---KPSLLKAALIAGATDGLG-IGYP 301
Db 237 DYTAAPGTSMTAPVHAGIAALLQ---AHPSTWTPDKVTALITADIIVKPEIADIAY- 292
Qy 302 SGNGQWGRVTLDKSL---NVAFVNETSSLTNQKATYSFTAQSGKPLKISLWSDAPAST 358
Db 293 ---GAGRVNVAIKYDDYAKLTFTGSDVADKGSATHTPDVSGATFVTATLYWDNAN--- 345
Qy 359 SASVTLNLDLVTAPNGTKYVGNDFTAIPYDNNWDRNNVNFVFINAPQSGTYTV 418
Db 346 -----SDLLYLYDPNGNQ-VDSYATAY-----GFEKVGYNPTDGTWIKVWS 389
Qy 419 YNVPPQGPQAPSLAIVN 434
Db 390 YS---GSANYQVDVWS 402

RESULT 13
US-10-090-624-4
; Sequence 4, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:

; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOCO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/10/090,624
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (428)..(428)
; OTHER INFORMATION: Xaa at position 428 is Gly or Val.
US-10-090-624-4

Query Match 18.0%; Score 404; DB 13; Length 522;
Best Local Similarity 30.0%; Pred. No. 1e-26;
Matches 131; Conservative 63; Mismatches 168; Indels 74; Gaps 16;

Qy 18 GLYGQGVAVADTGLDTRNDSSMHEAFRGKITAIYALGRNTNNAN-----DPNGHGTH 71
Db 22 GYDGSITIGIITGID-----ASHPDLOGKV-----IGWVDFVNGRSYPYDDHGHGTH 70
Qy 72 VAGSVLNGTNSN-----KGMAPQANLVFQSVN-----DSNGGLGGLPSNVSTLFSQAYSAGARI 125
Db 71 VASIAAGTGAASNGKYKGMAPGAKLAGIKVLGADGSGSISTIIKGVENAVDNKDKYGIKV 130
Qy 126 HTNSWGPVNGAYTTDSRNVDDYVRKNDMAVLFAAGNEGPNCGTISAPGTAKNAITVGAT 185
Db 131 INLSGSSOSSDGTALSOAVNAWDAGLVVVVVAAGNKGPNKYTIGSPAASKVITVGA- 189
Qy 186 ENLRPSFGSVADNINHVAFQSRGPTKDGRIKPDVMAPTFILSARSSILAPDSFWANHD 245
Db 190 -----VDKYDVITFSFSSRGPTADGRLKPEVAPGNWIIAARAS---GTSMGQPIN 236
Qy 246 SKYAYMGTSMTAPIVAGNVAQLREHFINKRGITP---KPSLLKAALTAGATDIG-LGYP 301
Db 237 DYTAAAGTSMATPHVAGIAALLQ---AHPSTPDVKVTALLETADIIVKPEIADIAY- 292
Qy 302 SGNGQGWRTLDKSL---NVAFVNETSSLTNOKATYSTAQSGLKPLKISLVMSDAPAST 358
Db 293 -----GAGRVNAYKAINVDYAKLVFTGYVANKGSQTHQFVIGASFVATLYWDNAN--- 345
Qy 359 SASVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGRNENNVFINAPQSGTYTVEVOA 418
Db 346 -----SDLDLVLDPNGNQ-VDYSYTAAY-----GFEKVGYNPTDGTWTIKVVS 389
Qy 419 YNVPGQPQAFSLAIVN 434
Db 390 YS---GSANYQVDVVS 402

RESULT 14
US-10-090-624-16
; Sequence 16, Application US/10090624
; Publication No. US20020123235A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE

; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/10/090,624
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-10-090-624-16

Query Match 18.0%; Score 404; DB 13; Length 654;
Best Local Similarity 30.0%; Pred. No. 1.4e-26;
Matches 131; Conservative 63; Mismatches 168; Indels 74; Gaps 16;

Qy 18 GLYGQGVAVADTGLDTRNDSSMHEAFRGKITAIYALGRNTNNAN-----DPNGHGTH 71
Db 154 GYDGSITIGIITGID-----ASHPDLOGKV-----IGWVDFVNGRSYPYDDHGHGTH 202
Qy 72 VAGSVLNGTNSN-----KGMAPQANLVFQSVN-----DSNGGLGGLPSNVSTLFSQAYSAGARI 125
Db 203 VASIAAGTGAASNGKYKGMAPGAKLAGIKVLGADGSGSISTIIKGVENAVDNKDKYGIKV 262
Qy 126 HTNSWGPVNGAYTTDSRNVDDYVRKNDMAVLFAAGNEGPNCGTISAPGTAKNAITVGAT 185
Db 263 INLSGSSOSSDGTALSOAVNAWDAGLVVVVVAAGNKGPNKYTIGSPAASKVITVGA- 321
Qy 186 ENLRPSFGSVADNINHVAFQSRGPTKDGRIKPDVMAPTFILSARSSILAPDSFWANHD 245
Db 322 -----VDKYDVITFSFSSRGPTADGRLKPEVAPGNWIIAARAS---GTSMGQPIN 368
Qy 246 SKYAYMGTSMTAPIVAGNVAQLREHFINKRGITP---KPSLLKAALTAGATDIG-LGYP 301
Db 369 DYTAAAGTSMATPHVAGIAALLQ---AHPSTPDVKVTALLETADIIVKPEIADIAY- 424
Qy 302 SGNGQGWRTLDKSL---NVAFVNETSSLTNOKATYSTAQSGLKPLKISLVMSDAPAST 358
Db 425 -----GAGRVNAYKAINVDYAKLVFTGYVANKGSQTHQFVIGASFVATLYWDNAN--- 477
Qy 359 SASVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGRNENNVFINAPQSGTYTVEVOA 418
Db 478 -----SDLDLVLDPNGNQ-VDYSYTAAY-----GFEKVGYNPTDGTWTIKVVS 521
Qy 419 YNVPGQPQAFSLAIVN 434
Db 522 YS---GSANYQVDVVS 534

RESULT 15
US-10-156-761-13251
; Sequence 13251, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HOSIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109

```
; SEQ ID NO 13251
; LENGTH: 1208
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13251

Query Match          16.1%; Score 362; DB 14; Length 1208;
Best Local Similarity 30.1%; Pred. No. 1.7e-22;
Matches 135; Conservative 62; Mismatches 186; Indels 66; Gaps 17;

QY 8 VKADVAQSS-----YGLYGGQGVVAVADTGLDTRNDSSMHAEFRGKITAIYALG 57
Db 186 VEADMAESNAQIGTRAANDAGLTGGVTVAVLDTGVDT-----THPDLAGRVSRSKSF 239

QY 58 RTNNANDPNRGHGHVAGSVLNGTSEN---KGMAPQANLVFQSVNDSNGGLGLPSNVST 113
Db 240 DGEVADRNRGHHVHTSTVGGSGAASDGTFRGVAPCATLAVGKVL-SDQGAGSESQIIAG 298

QY 114 LFSQAYSAGARIHTNSWGAPVNGAYTTD--SRNVDDYVRKNDMAVLFAAGNEGPNGGTIS 171
Db 299 MEWAARDVRARIVSMSLGS-TEASDGTDPMAEAVDTLSEETGALFVVAAGNTGA-PSSIG 356

QY 172 APTAKNAITVGATENLRPSFGSYADNINHVAQFSSRGPTK-DGRIKPDVMAPTFILSA 230
Db 357 SPGAADSALTUGA-----VDSSDRAAYFTSAGSRHGDNALKPDLAAPGVDTFAA 405

QY 231 RSLAPDSSFWANHDSKYAYMGGTSMATPIVAGNVAQLREHFIKNRGITPKPSLILKAA-- 288
Db 406 RSQLAFGTCY-----YTSMSGTSMATPHVAGVAALLAEQHPDWTGARLKDALMSTSEQ 458

QY 289 LIAGATDIGLYPSGNQGWG-RVTLDKSLNVAFVNETSLSLNQKATYSFTAQSGKPLKI 347
Db 459 LDASVYQLGAGRVSVDPDVGARVATGSAADLGFHRPHDADRPVTKVTYNSSDTTVEL 518

QY 348 SLVWSDAPASTSASVTLVNDLQVITAP-----NGTKYGVNDFTPYDNNWGRNNVNVF 403
Db 519 SLAVRGAPACVA---TLA---DTALTVPAGHTAATTVTGDSKAPV-----GETSGQIV- 566

QY 404 INAPQSGTYTVEQAYNVFQGPQAFSLAI 432
Db 567 ---ASAGGAPVAHTAFGLVKEAERYSLTV 592

Search completed: March 31, 2004, 16:34:02
Job time : 34.2177 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:38 ; Search time 11.3508 Seconds
(without alignments) 3677.911 Million cell updates/sec

Title: US-09-985-689A-6
 Perfect score: 2250
 Sequence: 1 NDVARGIVKADVAQSSVCLY.....EVOAYNVPGQPQAFSLAIYN 434

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

Database : PIR 78:*

```

1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | | DB | ID | Description |
|------------|-------|-------|--------|----|--------|---------------------|
| | | Match | Length | | | |
| 1 | 518.5 | 23.0 | 1743 | 2 | T19279 | multidrug resistan |
| 2 | 495 | 23.0 | 1905 | 2 | T19267 | multidrug resistan |
| 3 | 344.5 | 15.3 | 444 | 2 | B83891 | intracellular alka |
| 4 | 324 | 14.2 | 442 | 2 | A69587 | intracellular alka |
| 5 | 318.5 | 14.2 | 806 | 2 | A41341 | microbial serine p |
| 6 | 314.5 | 14.0 | 138 | 2 | T21559 | pyrolysins (EC 3.4. |
| 7 | 291 | 12.9 | 799 | 2 | G83753 | subtilisin-type pr |
| 8 | 279 | 12.4 | 715 | 2 | JC4908 | alkaline serine pr |
| 9 | 274 | 12.2 | 580 | 2 | S11890 | serine proteinase |
| 10 | 273 | 12.1 | 1345 | 2 | T29090 | surface layer-asso |
| 11 | 272 | 12.1 | 419 | 1 | S25835 | subtilisin (EC 3.4 |
| 12 | 272 | 12.1 | 420 | 1 | S23407 | subtilisin (EC 3.4 |
| 13 | 270.5 | 12.0 | 757 | 2 | C84120 | subtilisin-type pr |
| 14 | 264.5 | 11.8 | 401 | 2 | I39574 | serine proteinase |
| 15 | 260.5 | 11.6 | 533 | 1 | A35742 | aqualysin (EC 3.4. |
| 16 | 259 | 11.5 | 894 | 2 | F69730 | cell wall-associat |
| 17 | 257 | 11.4 | 380 | 2 | A49778 | high-alkaline seri |
| 18 | 257 | 11.4 | 1331 | 2 | A72647 | probable surface l |
| 19 | 255 | 11.3 | 627 | 2 | D75393 | serine proteinase, |
| 20 | 254 | 11.3 | 519 | 2 | S71451 | haemolysin R4 (EC 3 |
| 21 | 252.5 | 11.2 | 534 | 1 | JS0173 | alkaline proteinas |
| 22 | 250.5 | 11.1 | 379 | 1 | SUBSL1 | subtilisin (EC 3.4 |
| 23 | 247 | 11.0 | 382 | 2 | I39780 | subtilisin (EC 3.4 |
| 24 | 245 | 10.9 | 321 | 1 | S27501 | alkaline proteinas |
| 25 | 243.5 | 10.8 | 910 | 2 | C69456 | subtilisin sendai |
| 26 | 243 | 10.8 | 488 | 2 | A11930 | proteinase import |
| 27 | 242.5 | 10.8 | 381 | 2 | JH0778 | subtilisin (EC 3.4 |
| 28 | 242.5 | 10.8 | 382 | 1 | SUBSN | subtilisin (EC 3.4 |
| 29 | 240.5 | 10.7 | 274 | 1 | SUBSD | subtilisin (EC 3.4 |

| | | | | | | |
|----|-------|------|------|---|--------|---|
| 30 | 240.5 | 10.7 | 381 | 1 | SUBSI | subtilisin (EC 3.4.21.1) |
| 31 | 240 | 10.7 | 378 | 2 | A33973 | high-alkaline seri |
| 32 | 239 | 10.6 | 272 | 2 | A32624 | subtilisin (EC 3.4.21.1) |
| 33 | 233 | 10.6 | 1167 | 1 | A35066 | streptococcal CSA |
| 34 | 238.5 | 10.6 | 381 | 1 | SUBSS | subtilisin (EC 3.4.21.1) |
| 35 | 238.5 | 10.6 | 381 | 2 | JQ1487 | subtilisin (EC 3.4.21.1) |
| 36 | 238 | 10.6 | 384 | 2 | JC4802 | alkaline proteinase |
| 37 | 235.5 | 10.5 | 769 | 2 | D86335 | alkaline proteinase - T2OH2.6 protein - |
| 38 | 235 | 10.4 | 402 | 1 | JU0332 | alkaline proteinase |
| 39 | 233.5 | 10.4 | 530 | 2 | A42605 | halolysin (EC 3.4.21.1) |
| 40 | 232 | 10.3 | 374 | 2 | I3781 | subtilisin (EC 3.4.21.1) |
| 41 | 232 | 10.3 | 1374 | 2 | D2593 | hypothetical prote |
| 42 | 232 | 10.3 | 1433 | 1 | A36734 | bacilloleptidase F |
| 43 | 231.5 | 10.3 | 275 | 2 | JC1085 | subtilisin (EC 3.4.21.1) |
| 44 | 231 | 10.3 | 613 | 2 | S75976 | hypothetical prote |
| 45 | 230.5 | 10.2 | 401 | 2 | A57690 | aerolysin precursor |

ALIGNMENTS

RESULT 1
T18279
multidrug resistance transport protein - slime mold (Dictyostelium discoideum)
C/Species: Dictyostelium discoideum
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T18279
R/Shaulsky, G.; Loomis, W.F.
submitted to the EMBL Data Library, June 1996
A/Reference number: Z18855
A/Accession: T18279
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1743 <SHA>
A/Cross-references: EMBL:U60086; NID:gl399914; PID:gl399915; PIDN:RAB03331.1
C/Genetics:
A/Gene: tagC

Query Match 23.0%; Score 518.5; DB 2; Length 1743;
Best Local Similarity 28.0%; Pred. No. 6.3e-25;
Matches 165; Conservative 77; Mismatches 155; Indels 193;

```

19  LYGGQVVAVADTGLDTR--NDS-----SMHEAFROKITAIALGRTNANDNGH 68
314 LRKGQITSLADTGLDGHCFDFSSKYFIPILNSVNLNHR-KVVTYITTTSDDDSKVDGH 372
69  GTHVAGSVLG-----NGTSNKGMAQANLVFQSVMDSNGLGGL--PSNVSTLFSQAY 119
373 GTHICGSAAGTPEDSSVNISSFGLATDKAIP--FDLASGSSSLTFPSPDLKQIYQPLY 429
120 SAGARIHTNSGA----PVNGAYTTDSRNVDDYVRKN-DMAVLFAAGNEPNGGTIS--A 172
430 DAGARVCHDSWGSVSVEGTYGSYSDTASIDDFLPTHDPFIILRAAGN---NEQVLSLLT 486
173 PGRAKNAITVGATENLR-----PSFGSYADNI----- 199
487 QSTAKNVTIVGAHQTHENYLTDPGNINYQSSVDINQELICDPDSRYCNTTACQULES 546
200 -----NHVAQFSRRGFTKDGRIKPDVMAFGTIL 228
547 NATTGLASCCPTLLRKSVIDAANTQPLLXNENNICSFSSKGPTHDRGMKPALVAPGEVIT 606
229 SARSLA-----PDSGFWANHDSKYAMGTSVATPIVAGNVAQLREH-----F 272
607 SARSENGANTTQCQDGSLL-PNTNALLA-IGTSMATSFAAAATTLRQYLVGDGYPTGSI 664
273 IKRNGITPKSLLKAAILIAGA-----TDIGLYPSGN-----QGWGRVT 311
665 VESNKLQGTSLLKALMINNAQLNGTFLQITSSITYPNSQVFNFAGASIVQWGAIR 724
312 LDKSLNVAFVNETS-----SUSTNOKATVSTF-- 338
725 MSNWLHVYNNNSNNNNKTSDGITKFDGIGGLDLRLVKPNQWKBEESLTGQNTSYCFYK 784

```

339 QY -----AQSCKPLK-----ISLWSDPASTSASVTILNDLAVITAPNGTKXVGNDFAPVD 390

785 Db PSSSSNSGNPIPRVVALTWITDPSFVAGAFNVNNDLDTM-----IYRDNGSTIFIS 839

391 QY N-----NWDGRNVENVFINAPQSCITYTVEVQAYNVPGQPAFS 429

840 Db NQGGSFCLAPGTOTLNNVEGIVHNPEPTMYRFMVAGTNVMPGPNQF 889

RESULT 2

T18267
 multidrug resistance protein - slime mold (Dictyostelium discoideum)
 C/Species: Dictyostelium discoideum
 C/Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T18267
 R/Shaulsky, G.; Kuspa, A.; Loomis, W.F.
 submitted to the EMBL Data Library, January 1995
 A/Description: An MDR transporter/serine protease gene is required for prestalk specialization
 A/Reference number: Z18950
 A/Accession: T18267
 A/Status: preliminary; translated from GB/EMBL/DBBJ
 A/Molecule type: DNA
 A/Residues: 1-1905 <SHA>
 A/Cross-references: EMBL:U20433; MID:G664839; PID:G664840; PIDN:AAA62212.1
 C/Genetics:
 A/Gene: taqB

| Query Match | Score | DB 2 | Length |
|-----------------------|-------|-------------------|--------|
| Best Local Similarity | 22.0% | 495 | 1905 |
| Matches 163 | 28.2% | Pred. No. 2.28-23 | |
| Conservative | 72 | Mismatches | 162 |
| Indels | 180 | Gaps | 23 |

```
QY 19 LYGGQVAVADTGLDGTGR--NDS-----SMHEAFPGKIITAIYALGRTNNANDP--N 66
DB 376 LRFGQILSIADTGLDGSCHFFSSKYPIPFNQVNNENHRKVVTVI-----TYHNDNVYN 430
```

| | | | |
|----|-----|--|-----|
| Qy | 67 | GHGTHVAGSVLNG-----TSNKGMAFOANLVFQSVVDSNCGGLGGLPSNVSTLFSQAY | 119 |
| Dh | 431 | GHGTHCVGSAAGTPEDSSWAISPSFGLATDAKIAFYD-LSSGSGSEPTPEDYSOMYKPLY | 489 |

[illegible][illegible][illegible]

```
QY      233 S-----LAPDSSFANHDSKVAYMGGTSMATPIVAGNVAQLREHF-----IXNR 276
```

```
QY      277 GITYPSELLKAAIAGATDIGLY-----PSGN-----OQGRVTLDKSLN 317
```

```

00 318 VAFVNET-----SSLSNQKAT-----YSFTAQ-----SGKP 344
74 318 VAFVNET-----SSLSNQKAT-----YSFTAQ-----SGKP 344

```

[illegible]

| | | | |
|----|-----|--|-----|
| DD | 847 | KIVATLVWTDPPPSYSGAKFLVNNDLLLLLNSDDUUSIIIGNSSGSLQFAGKVAQF---- | 90 |
| QY | 393 | WDGRNNVENVFINAPOSGTITVEVOAYNVPOGPQAIFS | 429 |

DB 903 -DTLNNVEGIIINPTKAMNYKFTIAGTNVPIGPQKFS 938

RESULT 3

B83891
intracellular alkaline serine proteinase aprX [imported] - *Bacillus halodurans* (strain
C;Species: *Bacillus halodurans*
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: B83891
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir-
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A;Reference number: AB3650; MUID:20512582; PMID:11058132
A;Accession: B83891
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-444 <STO>
A;Cross-references: GB:AP001513; GB:BA000004; NID:G10174345; PIDN:BAB05649.1; GSPDB:GNO
A;Experimental source: strain C-125
C;Genetics:
A;Gene: aprX

Query Match 15.3%; Score 344.5; DB 2; Length 444;
Best Local Similarity 31.8%; Pred. No. 1.2e-14;
Matches 108; Conservative 45; Mismatches 98; Indels 89; Gaps 15;

| | | | |
|----|-----|---|-----|
| QY | 11 | DVAQSSVGLYGGQGVAVADTGLDTGRNDSMHEAFPKKITA-IYALGRTNNDPNC | 69 |
| | | | |
| | | | |
| Dh | 135 | EVTRNGEFTLTKDQWTVIAIDTGI-----YPHEDLEGRIKAFVDPNQREEPYDDNGH | 187 |
| | | | |
| | | | |

QY 70 THVAGSVLNGTGN-----KGMAPQANLVFQSVNDSNGLGGL-----P 108

198 THVAGGAGNGASAGGQVGGPAPNAVTVGVKLNNKQMGSLFESIMOGVWCIQYNEBHP 246

| | | | |
|----|-----|--|-----|
| QY | 109 | SNVSTLFSQAYSAGA-----RIHTNSWGAPVNGAYTTDSRNVDVYRKNDMAV | 156 |
| ST | 047 | CDRIHTNSGICCAATPYNEGCEGDMCTPIVETRNAGCI-----TV | 286 |

157 LFAAGNEGNGGTISAPGTAKNAITVGATENLRPSFGSYADINHNVAQFSRRGPTKGR I 216

QY 217 KPDWAPGTFILSARSSLAPDSSP-----WANHDSKYAYMGGSMTATPVGNAQLREH 271

QY 272 FIKRGITPKPSILKAALIAGTD-----IGLGYPES 302

RESULT 4

RESULT 4

A69587
intracellular alkaline serine proteinase aprX - Bacillus subtilis
C.Species: Bacillus subtilis
C.Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C.Accession: A69587
R.Kunst, F.; Ogasawara, N.; Mosser, I.; Allertini, A.M.; Alloni, G.; Azavedo, V.; Bertero,
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chalkley,
A.; Ehrlich, S.D.; Emerson, P.T.; Enian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Gallardo,
Nature 390, 249-256, 1997
A.Authors: Coulgier, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funai, S.; Galizzi, A.; Gallardo,
Fischer, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hulio, M.F.; Im-
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinios,
A.Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueeell,
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Ross, M.; Sadale, Y.; Sato, T.; Scanlon,
A.Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serca,
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpatra, P.; Tokutomi, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yanane, K.; Yasumoto, K.; Yoshida, K.; Yoshida,
A.Authors: Yoshikawa, H.F.; Zumbate, E.; Yoshikawa, H.; Yashiki, A.
A.Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A.Reference number: A69580; PMID:98044033; PMID:9384377
A.Accession: A69587
A.Status: preliminary; nucleic acid sequence not shown; translation not shown
A.Molecule type: DNA

A;Residues: 1-442 <KUN>
A;Cross-references: GB:Z99113; GB:AL009126; NID:q2634090; PIDN:CAB13610.1; PID:e1183385;
A;Experimental source: strain 168
C;Genetics:
A;Gene: aprX
C;Superfamily: subtilisin homology
F:146-398/Domain: subtilisin homology <SBT>
Query Match 14.4%; Score 324; DB 2; Length 442;
Best Local Similarity 31.6%; Pred. No. 2.4e-13;
Matches 106; Conservative 49; Mismatches 120; Indels 60; Gaps 13;
QY 11 DVAQSSYGLYGGQVAVADTGLDGRNDSSVHEAFKGTITAIYALGRTNAN-----D 64
DB 136 EYVRNGQTLTGKGVAVVDTGI-----YHPDLEGR-----IGFADMVNQKTEPYD 183
QY 65 PNGHGTTHVAGSVLNGTNS---KGMAPQANLVFQSVMSDGLGLPSNVSTLFSQAYS 120
DB 184 DNGHGTTHVAGSVLNGTNS---KGMAPQANLVFQSVMSDGLGLPSNVSTLFSQAYS 120
QY 121 AGARIHTNSWGPVNV-----GAYTDSRVVDYVRKNDMA-----VLFAGNENGP 166
DB 238 WCIOYNEADNPDEPIDTMSLGLDALRYDHEQEDPLVRAVEAWAGIIVCVAAAGNSGD 297
QY 167 GTTISAPGAKNAITVGTENLRPSFGSVADININVAOFSSRPGTKDGRKIPDVMAPTG 226
DB 298 SQTIASPGVSEKVTIVGALDDNTA-----SSDDTVASFSSRPGTPTVYKPKDILAPGV 353
QY 227 ILSARSAPDSF-----WANHDSKYAYMGTSMTPTVAGNVAGLREHFKNRGTIPK 281
DB 354 IISLRS---PNSYIDKLKSSRVGSOYFTMSGTSMATPICAGIAALILQ---QNPDLTPD 407
QY 282 PSLKKAALTAGTDIGLVPSCGQNGRVTLDKSL 316
DB 408 E---VKELKNGTKDKWEDDPNIYAGAVNAENSV 439
RESULT 5
A41341
microbial serine proteinase (EC 3.4.21.-), minor (vpr), precursor - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 17-Jul-1992 #sequence revision 17-Jul-1992 #text_change 20-Jun-2000
A;Accession: A41341; S39700; D69730
R;Sloma, A.; Ruffo Jr., G.A.; Theriault, K.A.; Dwyer, M.; Wilson, S.W.; Pero, J.
J. Bacteriol. 173, 6889-6895, 1991
A;Title: Cloning and characterization of the gene for an additional extracellular serine
A;Reference number: A41341; MUID:92041574; PMID:1938892
A;Accession: A41341
A;Molecule type: DNA
A;Residues: 1-806 <SLO>
A;Cross-references: GB:M76590; NID:G143819; PIDN:AAA22881.1; PID:G143820
A;Accession: B41341
A;Molecule type: protein
A;Residues: 161-195 <SL2>
R;Glaser, P.; Kunst, F.; Arnaut, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu,
A.; Rapoport, G.; Danchin, A.
Mol. Microbiol. 10, 371-384, 1993
A;Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region fr
A;Reference number: S39655; MUID:95020537; PMID:7934828
A;Accession: S39700
A;Molecule type: DNA
A;Residues: 1-806 <GLA>
A;Cross-references: EMBL:X73124; NID:G413923; PIDN:CAA51601.1; PID:9580871
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Poulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, M.; Kurita, K.; Lapilus, A.; Lardinois,
A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: D69730
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-806 <KUN>
A;Cross-references: GB:Z99123; GB:AL009126; NID:G2636240; PIDN:CAB15835.1; PID:G2636344
A;Experimental source: strain 168
A;Comment: The amino terminal sequence of the mature protein and a molecular weight of
C;Genetics:
A;Gene: vpr
A;Start codon: TTG
C;Superfamily: microbial serine proteinase vpr; subtilisin homology
C;Keywords: hydrolase; serine proteinase
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-160/Domain: propeptide #status predicted <PRO>
F:180-548/Domain: subtilisin homology #status atypical <SBT>
Query Match 14.2%; Score 318.5; DB 2; Length 806;
Best Local Similarity 23.9%; Pred. No. 1.2e-12;
Matches 132; Conservative 53; Mismatches 155; Indels 213; Gaps 19;
QY 18 GLYGGQVAVADTGLDGR-----NDSSMHEAFKGTITAIYALGRTN 60
DB 177 GYTGKGIKVAITDGVNHPDLKQFGYKGVDFVNDVDPKPTTG----- 224
QY 61 NANDPNC---HGTHVAGSVLNGTNSKGMAPQANLVFQSVMSDGLGLPSNVSTLFS 116
DB 225 --DPRGEATDHGTHVAGTVAANGTI--KGVAPDATLLAYRVLGPGG--SCTTENVIAGVE 278
QY 117 QAYSAGARIHTNSWGPVNGAYTDSRVVDYVRKNDMAVLFAAGNENGPNTISAPGTA 176
DB 279 RAVQDGADVMNLSLGNLNPNPDWATSTAL--DWAMSEGVAVTNGNSGPNGTWVGSGFTS 337
QY 177 KNATVIGATE---NLRFSPFGSY----- 195
DB 338 REALSVGATQLPNEYAVTFGSSYSSAKVMGYNKEDDVKALNNKEVELVEAGIGEAQDFEG 397
QY 196 -----ADNI----- 199
DB 398 KDLTKGVAVVKRGSIAFVADKADNAKAGAGIMVYNNLSGEIANVFGMSVPTIKLSLED 457
QY 200 -----NHVAQFSSRGPTKD--GRIPDVMAPGTFILSARS 232
DB 458 GEKLVSAKAGETTTFKLTVSKALGQVADFSRGPWMDTWIKPDISAPGVNIVSTIP 517
QY 233 SLAPDSSFWANHDSKYAYMGTSMTPTVAGNVAGLREHFKNRGTIPKPSL--LKAALI 290
DB 518 THDPD-----HPYGYGSKQGTSMASPHIAGAVAIKQ-----AKPKWSVEQIKAAIM 564
QY 291 AGATDI---GLGYPSGNGQGRVTLDKSLNVAFVNETSSLSTNQKATYGTAGSGKPLK 346
DB 565 NTAVTLKSDQGVVPHNAAGSARI---MNA--IKADSLVSPGSIYGYFLKENGNETK 619
QY 347 ISLVMSDAPASTASVTLVNDLDDLVITAPNGTKYGVNDFTAPYDNNWDG--RNNVENVFI 404
DB 620 -----NETFTIENQ-----SSIRKSYTLEYSFNGSGISTSGTSRVVI 656
QY 405 NAPSQGYTVTEVQ 417
DB 657 PAHOTGKATAKVK 669

RESULT 6

T28159
Pyrolysin (EC 3.4.-.-) - Pyrococcus furiosus
C;Species: Pyrococcus furiosus
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 24-Oct-2000
C;Accession: T28159

R;Voorthorst, W.G.B.; Eggen, R.I.L.; Geerling, A.C.M.; Platteeuw, C.; Siezen, R.J.; Vos, J. Biol. Chem. 271, 20426-20431, 1996
A:Title: Isolation and characterization of the hyperthermostable serine protease, pyroly
A:Reference number: Z20481; MUID:96355370; PMID:8702780
A:Accession: T23159
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1398 <VOO>
A:CROSS-references: EMBL:U55835; NID:g1556462; PID:g1556463; PIDN:AAB09761.1
A:Experimental source: DSM3638
C:Genetics:
A:Gene: p1s
C:Keywords: hydrolase; serine proteinase

Query Match 14.0%; Score 314.5; DB 2; Length 1399;
Best Local Similarity 24.1%; Pred. No. 4.6e-12;
Matches 126; Conservative 50; Mismatches 148; Indels 127; Gaps 15;
QY 21 GQGVAVADTGLDGRNDS-----SMHAFGKITAIALYALGRTNANDPN--- 66
DB 301 GNGYDIAYDVTDDYDFTDEVPLGQYNTYDVAVFSYYGPLYVLA-----EIDPNGE 354
QY 67 -----GHGTHVAGSVLNGTNS-----KG 85
DB 355 YAVFGWDGHHGTHVAGTVAGYDSNDAMDLSWYSGEWEVSRGLYGDYTNVTTDTVQG 414
QY 86 MAPQANLVFQSVNDSNGGLCLPSNTSLFSQAYSAGARIHTNSWG--APVNGAYTDSR 143
DB 415 VAPGAQIMAIRLRSQD--RGSMDIIEGTYAATHGADVINSGLGNAFYLDGTDPSV 472
QY 144 NVDDYVRKDMAVLFAAGNEGPGGTISAPGTAKNAITVGATENLRPSFGSYAD----- 197
DB 473 AVDELTEKYGVFVIAAGNEGPGINIVGSGVATKAITVGAA--VPINVGTVYSQALGYP 531
QY 198 -----NINHVAFSSRGPTKDGRIKPDVWAPGTFILSARSLAPDSSTWANHD 245
DB 532 DYTGFFYFPAYTNV--RIAFFSSRGPRIDGEIKENVVAPGYGYSLLPMWIGGADF----- 585
QY 246 SKYVMGTSWATPIVAGNVAQLREHFKNRGITPKPSLLKALIAACATDI-----GLG 299
DB 586 -----MSGTSWATPHVSGVALLISG--PKPEGIYVNPDIKKVLESATWLEGDPYTGQK 639
QY 300 YPSNQGWGRTVLDKSLNVAFNVTSLSTNQKATYSFTAQSKPLKISLWSDAPASTS 359
DB 640 YTELDOGHGLVNVTKSWEI-----LKAINGTTLPIVDHWADKSYSDF 681
QY 360 A---SVTLVNDLIDLITAPN-----GTKYVGN 383
DB 682 AEYLGVDVIRGLYARNISIPDIVWHIKYVGD 712

RESULT 7
G83753
subtilisin-type proteinase (EC 3.4.21.-) vpr precursor [similarity] - Bacillus halodurans
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: G83753
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: G83753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-799 <STO>
A:CROSS-references: GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BA04550.1; GSPDB:GN00
A:Experimental source: strain C-125
C:Genetics:
A:Gene: vpr
C:Superfamily: microbial serine proteinase vpr; subtilisin homology
C:Keywords: hydrolase; serine proteinase
F;1-29/Domain: signal sequence #status predicted <SIG>

Query Match 12.9%; Score 291; DB 2; Length 799;
Best Local Similarity 24.1%; Pred. No. 6.8e-11;
Matches 148; Conservative 57; Mismatches 165; Indels 244; Gaps 26;
QY 18 GLYGQGVAVADTGLDGRNDSMHAFGKITAIALYALGR-----TNN-----AN 63
DB 171 GYTGEGITVAILDTGVDYTHPD-----LVHAFGDKGWDIFDNDPDPQTTPG 218
QY 64 DPNG-----HGTHVAGSVLNGTNSKGMFAQANLVFQSVNDSNGGLPSNVSTL--FSQ 117
DB 219 DPRGIETHGTHVAGTVAANGLI--KGVAIPANLLAYRVL-----GPGRGSTAGVIAGIER 273
QY 118 AYSAGARIHTNSWAPVNGAYTTDSNVDDYVRKDMAVLFAAGNEGPGGTISAPGTAK 177
DB 274 AVQGDADIMLSLNTLNDPDPFATSTALDWMAGGVAVT--SNGNSGPNWTVGSPGTSR 332
QY 178 NAITVGATENLRPSFGSY-----ADNINH----- 201
DB 333 DAISVGAT--RLPYNKYKASVFTSGIDYPSADIMGFPDSDELLELDGETVEYAFAGLG 389
QY 202 ----- 201
DB 390 KPGDFEGVDVEGKIALIVRGEIPFVEKAENAKAAGAVGAILIYNNVAGVQPTVPLAIPTI 449
QY 202 -----VAQFSSRGPT--KDGRIKPDVWAPGTTF 226
DB 450 MLSNEDGLKMRNELENGQNTVFTSIEPDKLVGTVADFSRSGPVMHTWMIKPDVSAPGVA 509
QY 227 ILSARSLAPDSFHWANDSKYAVMGTSWATPIVAGNVAQLREHFKNRGITPKPSLLK 286
DB 510 IVSTPIPTHQDDPY-----GYGSRQGTSMASPHVAGAAALLLEAH--FNMGV-----DHVK 558
QY 287 AALIAGATDI-----GLGYPGNGOQ-----WGRVTLDKSLNVAF 320
DB 559 AALMNTAENLVDENGENYFPHNTQAGSIRIVDAISETLVTGSHSGFTTKERQKQVER 618
QY 321 VNETSSLSNQKATYSFTAQ--SGKP-----LKISLWSDAPASTSAVTLVNDLIDLITAP 375
DB 619 QHFTIHLNKRKTYQPDVQFAGNPDGIKVTKSKNLRVQPKTYQ--KINFNVQVDARKLDP 677
QY 376 ---NGTKYVGNDDFTAYDNNWNGSNVE---NVFINA---POSGTIVTEV-----QA 418
DB 678 GYEGTITVS-----DGSQTVETILFVSEPDYPRVTTPLDIDENGVLFGSA 726
QY 419 YNVPGQPAFSLAI 432
DB 727 Y-LPNGAEFGLWI 739
RESULT 8
JC4908
alkaline serine proteinase (EC 3.4.-.-) I precursor - Alteromonas sp.
N:Alternate names: subtilase
C:Species: Alteromonas sp.
C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 08-Oct-1999
C:Accession: JC4908
R;Fujisubo, H.; Miyamoto, K.; Tanaka, K.; Kaidzu, Y.; Imada, C.; Okami, Y.; Inamori, Y.
Biosci. Biotechnol. Biochem. 60, 1284-1288, 1996
A:Title: Cloning and sequence analysis of a protease-encoding gene from the marine bact
A:Reference number: JC4908; MUID:97141200; PMID:8987544
A:Accession: JC4908
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-715 <TSU>
A:CROSS-references: DBEJ:D38600; NID:g1536787; PIDN:BAA18912.1; PID:d1019647; PID:g2160
A:Experimental source: strain O-7
C:Comment: This enzyme belongs to class I subtilisin-like family. It is a chelator-sens
C:Genetics:
A:Gene: aprI
C:Superfamily: subtilisin homology
C:Keywords: hydrolase
F;1-40/Domain: signal sequence #status predicted <SIG>
F;41-150/Domain: amino-terminal propeptide #status predicted <ATP>

F;151-496/Product: alkaline serine protease I #status predicted <MAT>
F;182-452/Domain: subtilisin homology <SBT>
F;497-715/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;239-294,335-372,478-481/Disulfide bonds: #status predicted

Query Match 12.4%; Score 279; DB 2; Length 715;
Best Local Similarity 25.9%; Pred. No. 3.4e-10;
Matches 124; Conservative 51; Mismatches 168; Indels 136; Gaps 22;

QY 21 GQGQVAVADTG---LDTGRNDSSMHEAFRGKITAIYALGRTNNDP----- 65
DB 182 GQGQVAVADTG---LDTGRNDSSMHEAFRGKITAIYALGRTNNDP----- 65
QY 66 -----NGHGHVAG---SVLNGTSSNKGMAPQANLVFQSVMDNSGGGLGP 108
DB 242 DSSGQVPRADQSSWHGTHVAGTVAVTNGEGVAGVADAKVPRVL---GKCGGT 298
QY 109 SNVSTLFSQAYSAGARIHNSMGAPV-----NGAYTTDSNRVDYVAKNDMAVLFAAG 161
DB 299 SDIADGIWASGSDRVPANANPANNVNNSLGGGAGSATTQAINQARNNGTVIVIAAG 358
QY 162 NEGPNGGTTISAPCTAKNAITVGAT--ENLRPFGSYADNINHVAFQSSRGPTKDGRIKPD 219
DB 359 NDNDSANYN-PQNCNGVNVNVSVDGSGRAYSYNGANI-----D 398
QY 220 VMAPGTFILSARSSLAPDSSFWANHDS-----KYAYMGGTSMATPIVAGNVAQLR-- 269
DB 399 VAAAPG---CAQSFADDPGILLSTHNSGSGAPNSDYSYSGTSMAPHPHAGVAAALIKQA 454
QY 270 -----EHTFN--RGITKPSLLKAL--AGATDIGL-----YPSGNGQGRVTL 312
DB 455 KPSATPEVETILKNTTRSFAGSCNCGTGVDDAAVNEALGDVVTPGTN-----TL 508
QY 313 DKSINVAFVNETSLSSTNQKATSYFTAQSGKPLKISLWSDAPA-STASVTL---VNDL 368
DB 509 ED--GVAKTGLSAGASNQFFTF-----DVPAGKTNTVFTMSGTGDA 549
QY 369 DLVI---TAPNGKYVGNDFTPADYNNWGRNENNFVFNAPQSGTYTVEVOAYNVPOG 424
DB 550 DLYVKLGSPQTSYSSY---DCREYEG---GNAEVCSEFADPAQAGTYHYWINGYKAYS 599

RESULT 9
S11890
serine proteinase (EC 3.4.21.-) precursor, extracellular - Xanthomonas campestris pv. ca
N/Alternate names: subtilisin-related proteinase
C/Species: Xanthomonas campestris pv. campestris
C/Date: 21-Nov-1993 #sequence_revision 07-Feb-1997 #text_change 03-Dec-1999
C/Accession: S11890
R/Liu, Y.N.; Tang, J.L.; Clarke, B.R.; Dow, J.M.; Daniels, M.J.
Mol. Gen. Genet. 220, 433-440, 1990
A/Title: A multipurpose broad host range cloning vector and its use to characterise an e
A/Reference number: S11890; MUID:90251253; PMID:2187155
A/Accession: S11890
A/Molecule type: DNA
A/Residues: 1-580 <LTU>
A/Cross-references: EMBL:X51635; NID:948533; PIDN:CAA35962.1; PID:948534
A/Experimental source: Xanthomonas campestris pv. campestris
A/Note: the sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 205-Ala
C/Superfamily: subtilisin; subtilisin homology
C/Keywords: extracellular protein; hydrolase; serine proteinase
F;1-32/Domain: signal sequence #status predicted <SIG>
F;168-423/Domain: subtilisin homology <SBT>

Query Match 12.2%; Score 274; DB 2; Length 580;
Best Local Similarity 27.7%; Pred. No. 5.4e-10;
Matches 130; Conservative 53; Mismatches 157; Indels 130; Gaps 25;

QY 21 GQGQVAVADTG---LDTGRNDSSMHEAFRGKITAIYALGRTNNDP----- 64
DB 168 GSGTVVAVIDTGITSHADLANILAGYDFISDATTARDGNGRDSNADEGDVAAVECGA 227
QY 65 -----PNGHGHVAG---SVLNGTSSNKGMAPQANLVFQSVMDNSGG----- 103

DB 228 GIPAASSWHGTHVAGTVAAVTNTTGVAGTAYGAKVVPVRVLGKCGGSLSDIADAIWA 287
QY 104 ----LGLPSNV--STLFSQAYSAGARIHNSMGAPVNGAYTTDSNRVDYVAKNDMAVL 157
DB 288 SGGTVSGIPANANPAEIVNNLSLGGGSCSTTMQNA-INGAVSRGT-----TVV 334
QY 158 PAAGNEGPNGGTTISAPCTAKNAITVGATEN--LRPFGSYADNINHVAFQSSRGPTKDG 215
DB 335 VAAGNDASVSG--SLPANCANVIAVAATTSAGAKASYNFGTGI----- 377
QY 216 IKPDVMAPGTFILSARSS--LAPDSSFWANHDSKYAYMGGTSMATPIVAGNVAQLRHEFI 273
DB 378 ---DVSAPGSSILSTLNSGTTTFS-----ASYASYNSTSVASPHVAGVVALVQS--V 425
QY 274 KRGITPK--PSLLK--AALIAGATDIGLPGSQNGQWGRVTLKSLNVAFVNETS----- 325
DB 426 APTALTFAAVETLLKNTARALPGAC-----SGCGAGIWNADAATAA-INGSGGGG 477
QY 326 ---SLSTNQKATSYFTAQSGKPLKISLWSDAPASVTLV-----NOLDLVI---TA 374
DB 478 GGGNTLNGTPVTGLGATGAELNYTIT---VPAG--SGTLTVTTSGGSGDADLYVRAGSA 533
QY 375 PNGTKYVGNDFTPADYNNWGRNENNFVFNAPQSGTYTVEVOAYNVPOG 424
DB 534 PTDSAYT---CREYRS-----GNAETCTITAP--SGTYVYVRLKAYSTFSG 573

RESULT 10
T29090
surface layer-associated STABLE proteinase - Staphylothermus marinus
N/Alternate names: hyperthermostable proteinase
C/Species: Staphylothermus marinus
C/Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C/Accession: T29090
R/Mayr, J.; Lupas, A.; Kellermann, J.; Eckerskorn, C.; Baumeister, W.; Peters, J.
Curr. Biol. 6, 739-749, 1996
A/Title: A hyperthermostable protease of the subtilisin family bound to the surface lay
A/Reference number: Z20559; MUID:96385442; PMID:8793300
A/Accession: T29090
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-1345 <MAY>
A/Cross-references: EMBL:U57968; NID:gl374755; PID:gl374756; PIDN:AAB02323.1
A/Experimental source: strain F1
C/Function:
A/Description: probably serves an exodigestive function related to the organism's energ
A/Note: stoichiometric S-layer component

Query Match 12.1%; Score 273; DB 2; Length 1345;
Best Local Similarity 25.8%; Pred. No. 1.9e-09;
Matches 119; Conservative 68; Mismatches 154; Indels 120; Gaps 22;

QY 46 FRGKITAIYALGRTNNDPNGHGHVAG---GSYL-----GNGTSNK--GMAPOANLV 93
DB 445 YQGRYAL-----VSDFHGHGTSVATVIASGRVLYLDYGDGKLYRMGVAPGAKIA 496
QY 94 FQSVMDNSGGGLGPSNVSTLFSQAYSAG-----ARHTNSW 130
DB 497 -----GGDAWLGLNLVL--EAWLAGFNIVTEEDGYVYLSLDPFGPHRADIIISNW 546
QY 131 CA-----PVNGAYTTDSNRVDYVAKNDMAVLFAAGNEGPNGGTTISAPGTAK 177
DB 547 GSIYINFWLQFPDGIDYRSFMDLAIARNVLIQDHVTIVFAAGNEGPYSSNGCAPGTGL 606
QY 178 NAITVGATE--NLRPSPG---SYADNINHVAFQSSRGPTKDGRIKPDVMAPGTFILSARS 232
DB 607 LVITAGASTLWDYTRIYGYPEGYAD---EVIPESSRGPTGQYKPKPDIVNIGAFEWASTR 663
QY 233 SLAPDSGFWANHDSKYAYMGGTSMATPIVAGNVAQLRHEFIKRGITPKPSLLKAAALIAG 292
DB 664 TI--DORGYGAQPD-----VFGGTSEATPYTSGTLALVFQAYKEVYNTTDPVTKAILKSS 718

QY 293 ATDIGLYPSGNOGWGRVTLKSLNVAFFVNETSSLSTNQKATVSTFAQSGKPLKISLVMS 352
 DB 719 AKDI--WYPAFSGSGRVDKAAADTVFISE-----WLAIVSEGIQAFLENYIT 766
 QY 353 DAPASTASVTLVNDLDLV-ITAPNGTK-----YVGN-----DFTAPYDNN 392
 DB 767 DFGPYIGVLPYLADTDYGVVWPGSGSKNFTLIVGVNGAVSLSAMNTVLVKEYTV-YDGV 825
 QY 393 WDRGNVNFVINAPO---SGTVTE--VOAYNVPGQPQAF 428
 DB 826 YDSVG---LFLKVPKYVSGADYVEVVQLENWNTYPPGVF 863

RESULT 11
 S25835
 subtilisin (EC 3.4.21.62) precursor - Bacillus sp. (strain TA41)
 C:Species: Bacillus sp.
 C>Date: 22-Nov-1993 #sequence_revision 20-Feb-1995 #text_change 18-Jun-1999
 C:Accession: S25835
 R:Davall, S.; Feller, G.; Narinx, E.; Gerday, C.
 Gene 119, 143-144, 1992
 A:Title: Sequence of the subtilisin-encoding gene from an antarctic psychrotroph Bacillus
 A:Reference number: S25835; MUID:93012966; PMID:1398082
 A:Accession: S25835
 A:Molecule type: DNA
 A:Residues: 1-419 <DAV>
 A:Cross-references: EMBL:X63533; NID:g40198; PIDN:CAA45096.1; PID:g40199
 C:Superfamily: subtilisin; subtilisin homology
 C:Keywords: extracellular protein; hydrolase; serine proteinase
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-110/Domain: propeptide #status predicted <PRO>
 F:111-419/Product: microbial serine proteinase #status predicted <MAT>
 F:135-373/Domain: subtilisin homology <SBT>
 F:144,184,359/Active site: Asp, His, Ser #status predicted

Query Match 12.1%; Score 272; DB 1; Length 419;
 Best Local Similarity 31.7%; Pred. No. 4.6e-10;
 Matches 83; Conservative 36; Mismatches 105; Indels 38; Gaps 10;

QY 21 GQGVVAVADTGLDGRNDSMEAFRGKITAIALYAGRTNANDPNHGHTHVAGSLNG 80
 DB 135 GAGINAVLDTGVNTNHPDLIS--NNVEQCKDFTVGTNFTDNTDCTDRQGHGTHVAGSALANG 193
 QY 81 TSNK---GMAPQANLVFQSVMDNSGGLGFLPSNVSTLFSQAYSAGARIHTN-----SWGA 132
 DB 194 GTSGGVGVGAPEADLWAYKVLGDDG--SGVADDIAEAIHAGDQATALNTKVVINSLGS 251
 QY 133 PVNGAYTDSRVDDYVRKNDMAVLFAAGNEGPNGTISAPGTAKNAITVGTATENLRPSF 192
 DB 252 SGSSLIIT---NAVVDYAYDKGVLIIAAGNSGPKSGISGYPGALYNAVAALNTIQN- 307
 QY 193 GSYADNINHVAQFSSRGPTKDG-----RIKPDVMAPGTFFILSARSSLAPDSSFWANHDS 246
 DB 308 GTY-----RVADFSSRGHKRTAGDYVIQKGDVEISAPGAIVST-----W--FDG 350
 QY 247 KYAYMGTSNATPIVAGNVAQL 268
 DB 351 GYATISGTSNAPHAAGLAAKI 372

RESULT 12
 S23407
 subtilisin (EC 3.4.21.62) 1 precursor - Bacillus sp. (strain TA39)
 C:Species: Bacillus sp.
 C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 18-Jun-1999
 C:Accession: S23407
 R:Narinx, E.; Davail, S.; Feller, G.; Gerday, C.
 Biochim. Biophys. Acta 1131, 111-113, 1992
 A:Title: Nucleotide and derived amino acid sequence of the subtilisin from the antarctic
 A:Reference number: S23407; MUID:92256481; PMID:1581352
 A:Accession: S23407
 A:Molecule type: DNA
 A:Residues: 1-420 <NAR>

A:Cross-references: EMBL:X62369; NID:g40200; PIDN:CAA44227.1; PID:g40201
 C:Genetics:
 A:Gene: eubI
 C:Superfamily: subtilisin; subtilisin homology
 C:Keywords: extracellular protein; hydrolase; serine proteinase
 F:1-23/Domain: signal sequence #status predicted <SIG>
 F:24-111/Domain: propeptide #status predicted <PRO>
 F:112-420/Product: microbial serine proteinase #status predicted <MAT>
 F:136-374/Domain: subtilisin homology <SBT>
 F:145,185,360/Active site: Asp, His, Ser #status predicted

Query Match 12.1%; Score 272; DB 1; Length 420;
 Best Local Similarity 30.4%; Pred. No. 4.7e-10;
 Matches 96; Conservative 41; Mismatches 123; Indels 56; Gaps 15;

QY 21 GQGVVAVADTGLDGRNDSMEAFRGKITAIALYAGRTNANDPNHGHTHVAGS 75
 DB 136 GGGINAVLDTGVNTN-----HPDLNNVEQCKDFTVGTTYNNCTDRQGHGTHVAGS 189
 QY 76 VL---GNGTSNKGMAPOANLVFQSVMDNSGGLGFLPSNVSTLFSQAYSAGARIHTN----- 128
 DB 190 ALADGGTGVGVGAPADLWAYKVLGDDG--SGVADDIAAIAIRHAGDQATALNTKVVIN 247
 QY 129 -SMGAPVNGAYTDSRVDDYVRKNDMAVLFAAGNEGPNGTISAPGTAKNAITVGTATEN 187
 DB 248 MSLSGSSGESSLIITNAVN---YSYKGVLIITAAAGNSGYPGSGISGYPGALYNAVAALAN 304
 QY 188 LRPSFGSYADNINHVAQFSSRGPT-KDG-----RIKPDVMAPGTFFILSARSSLAPDSSFW 241
 DB 305 -KVENGTY-----RVADFSSRGYSWTDDGYAIQKGDVEISAPGAIVST-----W 348
 QY 242 ANHDSKYAYMGTSNATPIVAGNVAQLREHFINKGITPKPSLLKALAGATDGLGYP 301
 DB 349 --FDGGYATISGTSNAPHAAGLAAKIQAIPSAASNVDRGELQYRAY---ENDILSGYY 403

QY 302 SG-----NOQMGVRVTL 312
 DB 404 AGYGDFFASGFGPATV 419

RESULT 13
 C84120
 subtilisin-type proteinase (EC 3.4.21.-) BH3763 precursor [similarity] - Bacillus halod
 C:Species: Bacillus halodurans
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: C84120
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Saeaki, R.; Masui, N.; Fuji, F.; Hir
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: C84120
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-757 <STO>
 A:Cross-references: GB:AP001519; GB:BA000004; NID:g10176109; PIDN:BA07482.1; GSPDB:GNO
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: BH3763
 C:Superfamily: microbial serine proteinase vpr; subtilisin homology
 C:Keywords: hydrolase; serine proteinase
 F:1-20/Domain: signal sequence #status predicted <SIG>

Query Match 12.0%; Score 270.5; DB 2; Length 757;
 Best Local Similarity 24.2%; Pred. No. 1.3e-09;
 Matches 127; Conservative 48; Mismatches 159; Indels 191; Gaps 20;

QY 8 VKADVAQSSYGLVQGVAVADTGLDGRNDSMEAFRGKITAIALYAG 57
 DB 119 VRGMLDEGVHLTKGVKVAVIDTGIDYTHPD--LOSSYKGGYDFVDYDDDPMETIASQG 176
 QY 58 RTNNANDPNHGHTHVAGSLNGTSNKGMAPOANLVFQSVMDNSGGLGFLPSNVSTLFSQ 117
 DB 177 -----PPTLHGTHVSGIIAANG-QVKGVAPAEIYAYRALPGG--QGTTEQVIAIEK 227

RESULT 14
I39974
serine proteinase - Bacillus sp.
C:Species: Bacillus sp.
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 22-Jun-1999
C:Accession: I39974
R:Maciver, B.; McHale, R.H.; Saul, D.J.; Bergquist, P.L.
Appl. Environ. Microbiol. 60, 3981-3988, 1994
A:Title: Cloning and sequencing of a serine proteinase gene from a thermophilic Bacillus
A:Reference number: I39974; MUID:95085262; PMID:7993087
A:Accession: I39974
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-401 <RES>
A:Cross-references: GB:I29506; NID:9529979; PIDN:AAA63688.1; PID:9529980
C:Superfamily: subtilisin; subtilisin homology
F:151-361/Domain: subtilisin homology <SBT>

Query Match 11.8%; Score 264.5; DB 2; Length 401;
Best Local Similarity 32.6%; Pred. No. 1.3e-09;
Matches 86; Conservative 28; Mismatches 95; Indels 55; Gaps 11;

Qy 11 DVAQSSYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTNNDP---NG 67
Db 146 DVTKGS-----SQGEIADITGVD-----YHPDLGKVIKGYDF--VDNDYDMDLNN 192
Qy 68 HGVHVGAG---SVLGGTSNKGMAPOANLVFQSVMDNSGGLGGLPSNVSTLFSQAYSAGAR 124
Db 193 HGVHVGAGIAAETNNATGAGWPNTRILAVRALDRNG--SGLTSDIADAIYAADSGAE 250
Qy 125 IHTNSGAPVNGAYTTDSRNVDYVRKNDMAVLFAGNKGPNNGGTISAPGTAKNAITVGA 184
Db 251 VINLSLGC--DCHTTTLENNAVYAWKSGSVVVAAGNNG--SSITPEPASYENVIAVGA 305
Qy 185 TENLRPSFGSYADNHNHVAQPSRGTQDGRKIPDMAPGTFILSARSLAPDSSFWANH 244
Db 306 -----VDQDRLASFNNGTW-----VDVAPGVDIVSTIG----- 337
Qy 245 DSKYATMGTSMATPIVAGNVAQL 268
Db 338 -NEYAYMGTSMAHPVAGLALL 360

RESULT 15
A35742
aqualysin (EC 3.4.21.-) I precursor - Thermus aquaticus
C:Species: Thermus aquaticus
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: A35742; S00620; S00324
R:Terada, I.; Kwon, S.T.; Miyata, Y.; Matsuzawa, H.; Ohta, T.
J. Biol. Chem. 285, 6576-6581, 1990
A:Title: Cloning precursor structure of an extracellular protease, aqualysin I, with NH-
A:Reference number: A35742; MUID:90216674; PMID:2182621
A:Accession: A35742
A:Molecule type: DNA
A:Residues: 1-513 <TER>
A:Cross-references: GB:J90108; GB:D90108; GB:J05414; NID:9217171; PIDN:BA14135.1; PID:
A:Note: the authors translated the codon CTG for residue 470 as Val, and GGT for residu
R:Kwon, S.T.; Terada, I.; Matsuzawa, H.; Ohta, T.
Eur. J. Biochem. 173, 491-497, 1988
A:Title: Nucleotide sequence of the gene for aqualysin I (a thermophilic alkaline serin
A:Reference number: S00620; MUID:88225062; PMID:3286255
A:Accession: S00620
A:Molecule type: DNA
A:Residues: 75-442 <KWO>
A:Cross-references: EMBL:X07734; NID:948069; PIDN:CAA30559.1; PID:g602091
A:Note: part of this sequence, including the amino and carboxyl ends of the mature prot
R:Matsuzawa, H.; Tokugawa, K.; Hamaoki, M.; Mizoguchi, M.; Taguchi, H.; Terada, I.; Kwo
Eur. J. Biochem. 171, 441-447, 1988
A:Title: Purification and characterization of aqualysin I (a thermophilic alkaline seri
A:Reference number: S00324; MUID:88151937; PMID:3162211
A:Accession: S00324
A:Molecule type: protein
A:Residues: 128-170 <MATS>
C:Superfamily: subtilisin; subtilisin homology
C:Keywords: extracellular protein; hydrolase; serine proteinase
F:1-14/Domain: signal sequence #status predicted <SIG>
F:11-127/Domain: propeptide #status predicted <PRO>
F:118-408/Product: aqualysin I #status experimental <MAT>
F:157-364/Domain: subtilisin homology <SBT>
F:255-257/281-283/Region: Si specificity crevice #status predicted
F:409-513/Domain: carboxyl-terminal propeptide #status predicted <CPR>
F:166,197,349/Active site: Asp, His, Ser #status predicted

Query Match 11.6%; Score 260.5; DB 1; Length 513;
Best Local Similarity 28.2%; Pred. No. 3.3e-05;
Matches 122; Conservative 43; Mismatches 157; Indels 111; Gaps 23;

Qy 16 SYLGGQGVVAVADTGLDTRNDSSMHEAFRGKITAIY-ALGRTNNDPNGHGHVAG 74
Db 152 TYTATGGRVNVYVDTGIRT-----THREFGGRARVGYDALG--GNGQDCNCHGHVAG 203
Qy 75 SVLGGTSNKGMAPOANLVFQSVMDNSGGLGGLPSNVSTLFSQAYSAGARIHTNSWAP- 133
Db 204 TI---GGVTYGVAKAVNLVAVRVLDGNG--SGTSGV-----TAGVDWVTRNHRRA 250
Qy 134 -----VNGAYTTDSRNVDYVRKNDMAV---LFAAGNEGPNNGGTISAPGTAKNAITVGA 184
Db 251 VANWSLGGGVST---ALDNAV-KNSIAAGVYVAVAGNDNANACNYS--PARVAEALTVGA 305
Qy 185 T--ENLRPSFGSYADNHNHVAQPSRGTQDGRKIPDMAPGTFILSARSLAPDSSFWA 242
Db 306 TTSDDARASFNSGCV-----DLFAPGASIPSA-----WY 336
Qy 243 NHDSKYATMGTSMATPIVAGNVAQLREHFKNRGTTPKPSLLKAALIAAGT-----DIG 297
Db 337 TSDATQTTLNGTSMATPHVAGVAAVLE---QNPSAT--PASVASAILNGATTGLSLGIG 391
Qy 298 LGYP-----SGNQGWRGVTLDKSLNVAFNVTSSLTNOKATYSFTAQSGKPLKLSL 349
Db 392 SGSPNRLLYSLSSGSGSTAPCTSCSYTGSLSGPGDYNFQPNGTYYYS-----PAGTHR 446
Qy 350 VMSDAPASTASVTLVNDLDELVTAPNGTKY--VGNDFTPADYDNNWDGRNNVENVINAP 407
Db 447 AWLRGPAGT-----DFDLYLWRWDGSRWLTIVGSS--TGP-----TSEESLUSYG 488

QY 408 QSGTYTVEVCAYN 420
: | : | :
Db 489 TAGYLLWRIYAYS 501

Search completed: March 31, 2004, 16:10:22
Job time : 13.3508 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 7.51154 Seconds
(without alignment)
3008.498 Million cell updates/sec

Title: US-09-985-689A-6

Perfect score: 2250

Sequence: 1 NDVARGIVKADVAQSSVGLY.....EVQAYNVPGQPFQAFSLAIVN 434

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------|---------------------|
| 1 | 518.5 | 23.0 | 1743 | 1 TAGC DICDI | Q33868 dictyosteli |
| 2 | 495 | 22.0 | 1905 | 1 TAGB DICDI | P54683 dictyosteli |
| 3 | 318.5 | 14.2 | 806 | 1 SUBV_BACSU | P39141 bacillus su |
| 4 | 316.5 | 14.1 | 1398 | 1 PLS_PFRFU | P72186 pyrococcus |
| 5 | 274 | 12.2 | 580 | 1 EXPR_XANCP | P33314 xanthomonas |
| 6 | 272 | 12.1 | 420 | 1 SUBT_BACSP | P28842 bacillus sp |
| 7 | 264.5 | 11.8 | 401 | 1 THES_BACSP | Q45670 bacillus sp |
| 8 | 260.5 | 11.6 | 513 | 1 AOL1_THEAQ | P08594 thermus aqu |
| 9 | 259 | 11.5 | 894 | 1 WPR_A_BACSU | P54423 bacillus su |
| 10 | 257 | 11.4 | 269 | 1 SUBS_BACLE | P26600 bacillus le |
| 11 | 257 | 11.4 | 380 | 1 ELYA_BACAO | P27693 bacillus al |
| 12 | 257 | 11.4 | 380 | 1 ELYA_BACAO | P41362 bacillus cl |
| 13 | 255 | 11.3 | 269 | 1 PRTM_BACSP | Q99405 bacillus sp |
| 14 | 252.5 | 11.2 | 534 | 1 PROA_VIBAL | P16588 vibrio algi |
| 15 | 250.5 | 11.1 | 379 | 1 SUBT_BACLI | P00780 bacillus li |
| 16 | 249 | 11.1 | 269 | 1 SUBB_BACLE | P29599 bacillus le |
| 17 | 245 | 10.9 | 321 | 1 ISP_BACCS | P39140 bacillus cl |
| 18 | 242.5 | 10.8 | 381 | 1 SUBN_BACNA | P58835 bacillus su |
| 19 | 242.5 | 10.8 | 382 | 1 SUBN_BACNA | P00782 bacillus am |
| 20 | 241.5 | 10.7 | 1181 | 1 SCAT_STRPY | P58099 streptococ |
| 21 | 240.5 | 10.7 | 274 | 1 SUBD_BACLI | P00781 bacillus li |
| 22 | 240.5 | 10.7 | 381 | 1 SUBT_BACSP | P04189 bacillus su |
| 23 | 240 | 10.7 | 378 | 1 ELYA_BACSP | P20724 bacillus sp |
| 24 | 239 | 10.6 | 275 | 1 SUBT_BACPU | P07518 bacillus pu |
| 25 | 239 | 10.6 | 1167 | 1 SCAL_STRPY | P15926 streptococ |
| 26 | 238.5 | 10.6 | 381 | 1 SUBT_BACSA | P00783 bacillus su |
| 27 | 238.5 | 10.6 | 381 | 1 SUBT_BACST | P29142 bacillus st |
| 28 | 235 | 10.4 | 402 | 1 ALP_CEPAC | P39118 cephalospor |
| 29 | 234 | 10.4 | 422 | 1 TKSU_PYRKO | P58502 pyrococcus |
| 30 | 233.5 | 10.4 | 530 | 1 HLY_HALL17 | P29143 halophilic |
| 31 | 232 | 10.3 | 1433 | 1 SUBF_BACSU | P16397 bacillus su |
| 32 | 230 | 10.2 | 279 | 1 THET_THREVU | P40072 thermocactin |
| 33 | 227.5 | 10.1 | 1052 | 1 MSIP_CRIGR | Q922a8 cricetus |

ALIGNMENTS

RESULT 1

| TAGC | DICDI | 10.1 | 1052 | 1 | MSIP_MOUSE | Q9wtz2 | mus musculus |
|------|--|-------|------|------|--------------|--------|--------------|
| AC | Q23868 | 227.5 | 10.1 | 1052 | 1 MSIP_RAT | Q9wtz3 | rattus norv |
| DT | 01-NOV-1997 (Rel. 35, Created) | 224.5 | 10.0 | 1052 | 1 MSIP_HUMAN | Q14703 | homo sapien |
| DT | 01-NOV-1997 (Rel. 35, Last sequence update) | 224 | 10.0 | 319 | 1 ISP1_BACSU | P11018 | bacillus su |
| DT | 15-MAR-2004 (Rel. 43, Last annotation update) | 220.5 | 9.8 | 326 | 1 ISP_PAEPO | P29139 | paenibacill |
| DE | Prestalk-specific protein tagC precursor (EC 3.4.21.-) | 220.5 | 9.8 | 404 | 1 SMTI_WAGPO | Q9V778 | magnaporthe |
| GN | TAGC | 216.5 | 9.6 | 361 | 1 ELYA_BACHD | P41363 | bacillus ba |
| OS | Dictyostelium discoideum (Slime mold). | 215.5 | 9.6 | 1374 | 1 YQSG_CAEEL | Q09541 | caenorhabdi |
| OC | Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium. | 215 | 9.6 | 645 | 1 SUBE_BACSU | P16396 | bacillus su |
| OX | NCBI_TaxID=44689; | 215 | 9.6 | 1902 | 1 P2P_LACPA | Q02470 | lactobacill |
| RN | [1] | 214 | 9.5 | 293 | 1 PRTT_TRIAL | P20015 | tritirachiu |
| RP | SEQUENCE FROM N.A. | 213.5 | 9.5 | 1902 | 1 P1P_LACLC | P16271 | lactococcus |
| RC | STRAIN=AX4; | | | | | | |
| RX | MEDLINE=97140317; PubMed=8986798; | | | | | | |
| RA | Shaulsky G., Escalante R., Loomis W.F.; | | | | | | |
| RT | "Developmental signal transduction pathways uncovered by genetic | | | | | | |
| RT | suppressors."; | | | | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 93:15260-15265(1996). | | | | | | |
| CC | -1- FUNCTION: Intercellular communication via tagC may mediate | | | | | | |
| CC | integration of cellular differentiation with morphogenesis (By | | | | | | |
| CC | similarity). | | | | | | |
| CC | -1- SIMILARITY: In the N-terminal section; belongs to peptidase family | | | | | | |
| CC | S8. | | | | | | |
| CC | -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING | | | | | | |
| CC | TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY. | | | | | | |
| CC | -1- SIMILARITY: STRONG, TO TAGB. | | | | | | |
| CC | ----- | | | | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | | | | |
| CC | or send an email to license@isb-sib.ch). | | | | | | |
| CC | ----- | | | | | | |
| DR | EMBL; U60086; AAB0331.1; - | | | | | | |
| DR | PIR; T18279; T18279 | | | | | | |
| DR | DictyBase; DDB0001795; tagC | | | | | | |
| DR | InterPro; IPR003593; AAA ATPase. | | | | | | |
| DR | InterPro; IPR001140; ABC_TM_transp. | | | | | | |
| DR | InterPro; IPR003439; ABC_transporter. | | | | | | |
| DR | InterPro; IPR000209; Peptidase S8. | | | | | | |
| DR | Pfam; PF00664; ABC_membrane; 1 | | | | | | |
| DR | Pfam; PF00005; ABC_tran; 1 | | | | | | |
| DR | Pfam; PF00082; Peptidase S8; 1 | | | | | | |
| DR | PRINTS; PR00723; SUBTILISIN. | | | | | | |
| DR | SMART; SM00382; AAA; 1 | | | | | | |
| DR | PROSITE; PS00929; ABC_TM1F; 1 | | | | | | |
| DR | PROSITE; PS00211; ABC_TRANSPORTER_1; 1 | | | | | | |
| DR | PROSITE; PS00893; ABC_TRANSPORTER_2; 1 | | | | | | |
| DR | PROSITE; PS00136; SUBTILASE ASP; FALSE_NEG. | | | | | | |
| DR | PROSITE; PS00137; SUBTILASE_HIS; 1 | | | | | | |
| DR | PROSITE; PS00138; SUBTILASE_SER; FALSE_NEG. | | | | | | |
| KW | Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane; | | | | | | |

KW Signal.
 FT CHAIN 1 27
 FT CHAIN 28 1743
 FT CHAIN 316 642
 FT CHAIN 1450 1687
 FT CHAIN 962 982
 FT CHAIN 1027 1047
 FT CHAIN 1072 1092
 FT CHAIN 1157 1177
 FT CHAIN 1260 1280
 FT CHAIN 1288 1308
 FT CHAIN 325 325
 FT CHAIN 372 372
 FT CHAIN 637 637
 FT CHAIN 1485 1492
 FT CHAIN 42 46
 FT CHAIN 94 103
 FT CHAIN 643 646
 FT CHAIN 733 741
 FT CHAIN 786 792
 FT CHAIN 1337 1340
 FT CHAIN 1346 1352
 FT CHAIN 1353 1357
 FT CHAIN 1358 1364
 FT CHAIN 1381 1386
 FT CHAIN 1707 1729
 FT CHAIN 390 390
 FT CHAIN 536 536
 FT CHAIN 547 547
 FT CHAIN 614 614
 FT CHAIN 689 689
 FT CHAIN 735 735
 FT CHAIN 741 741
 FT CHAIN 776 776
 FT CHAIN 832 832
 FT CHAIN 887 887
 FT CHAIN 1251 1251
 FT CHAIN 1385 1385
 FT CHAIN 1386 1386
 FT CHAIN 1454 1454
 FT CHAIN 1704 1704
 FT CHAIN 1743 1743
 FT CHAIN 194145 MW; 12DB3632F729839 CRC64;
 SEQUENCE 1743 AA; 194145 MW; 12DB3632F729839 CRC64;
 Query Match 23.0%; Score 518.5; DB 1; Length 1743;
 Best Local Similarity 28.0%; Pred. No. 2.5e-25;
 Matches 165; Conservative 77; Mismatches 155; Indels 193; Gaps 23;
 QY 19 LYGGGVAVADTGLDGR---NDS-----SMHEAFSGKITAVALGRTNANDPXGH 68
 DB 314 LRKGGILSIADTGLDGHCFSDSKYPIPLNSVNLNR-KVTYITTSDDSDKVDGH 372
 QY 69 GTHVAGSVLG-----NNTSKGMAPOANLVFQSVMDNGLGLG--PSNVSTLFSQAY 119
 DB 373 GTHICGSAAGTPEDSSVNISSFSGLATDAKIAF---FDLASGSSSLTPPSDLKQLYQPLY 429
 QY 120 SAGARLHTNSWA-----PVNGAYTDSRVDDYVRKN-DMAVLFRAGNPGNGGIS--A 172
 DB 430 DAGARVCHDSWGSVVEGYTGSYSSTASIDFLFTHPFIILRAGN---NEQYLSLIT 486
 QY 173 PGAKNAITVGAENLR-----PFGSYADNI----- 199
 DB 487 QSTAKNVITVGAHOTIHENYLTDPGNVINYQSSVDINQELICDFDSRYCNVYTAQCLES 546
 QY 200 -----NHVAQFSSRGTKGRKIPKPDVWAGTFIL 228
 DB 547 NATGLASCCPTLRLKSVDAANTQPLLNNENNICFSKSGPTHDRMKPALVAPGEYIT 606
 QY 229 SARSSILA-----PDSSFWANHDSKVAYMGTSMATPIVAGNVAQLRHH-----F 272
 DB 607 SASNSGANTTDQCGDGL-ENTWALLA-IGTSWATSFAAAATTLRLQVLVDGYGPTGSI 664
 QY 273 IKRNGITPKSLILKAAALIAGA-----TDILGLYPSGN-----QNGRVT 311
 DB 273 IKRNGITPKSLILKAAALIAGA-----TDILGLYPSGN-----QNGRVT 311

665 VESNKLQPTGSLKALMINNAQLNGTFLQITSSITYPSNQVFENFAGASIVQWGAIR 724
 QY 312 LDKSLNVAFWNETS-----SISTNQKATYSFT-- 338
 DB 725 MSNWLHVNNNNNNNNKTSIDGITKFDGIGGLDLRLVKPNQWKEESLSTGQNTSYCTYK 784
 QY 339 -----AQSGKPLK---ISLWMSDAPASTSASVTLVNDLVLITAPNGTKYVGNDFAPYD 390
 DB 785 PSSSSSSNSGNNIPRVVATILVTPPSYAGAKFNLVNNLDLTM-----IYRONGSTIFYS 839
 QY 391 N-----NWDGRNNVNFVINAPOSCTVTVVQAVNVPOGPOAFS 429
 DB 840 NQGSSEFLGLAPTQDILNVEGIVNHTPTMTYRFVWAGTNVPMGPQNS 889
 RESULT 2
 TAGS DICDI STANDARD; PRT; 1905 AA.
 AC P54683;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Prestalk-specific protein tagB precursor (EC 3.4.21.-).
 GN TAGB.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 NCBI_TaxID=44689;
 [1]
 SEQUENCE FROM N.A.
 STRAIN=AX4;
 MEDLINE=95262903; PubMed=7744252;
 RA Shaulesky G., Kuspa A., Loomis W.F.;
 RT "A multidrug resistance transporter/serine protease gene is required
 for prestalk specialization in Dictyostelium.";
 RL Genes Dev. 9:1111-1122(1995).
 CC -!- FUNCTION: Inter-cellular communication via tagB may mediate
 integration of cellular differentiation with morphogenesis.
 CC -!- SIMILARITY: In the N-terminal section; belongs to peptidase family
 S8.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING
 TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.
 CC -!- SIMILARITY: STRONG, TO TAGC.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U20432; AAA62212.1; -.
 PIR; T18267; T18267.
 DR MEROPS; S08.0PM; -.
 DR DictyBase; DB0001964; tagB.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001140; ABC_TM_transp.
 DR InterPro; IPR003439; ABC_transporter.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF00664; ABC_membrane; 1.
 DR Pfam; PF00005; ABC_tran; 1.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PRODOM; PD000006; ABC_transporter; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS05029; ABC_TMIF; 1.
 DR PROSITE; PS0211; ABC_TRANSPORTER_1; 1.
 DR PROSITE; PS05093; ABC_TRANSPORTER_2; 1.
 DR PROSITE; PS0136; SUBTILASE_ASP; FALSE_NEG.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00139; SUBTILASE_SER; 1.
 KW Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane;
 Signal.

[illegible]

| | | | |
|----|-----|---|-----|
| Qy | 246 | SKYAYGGTSMATPIVAGNVQALREHPIKVRGTPKPSLLKAALIAATDI-----GLG | 299 |
| Db | 586 | -----MSGTSMATPHSVGVALLISG-AKAEGIYVNPDIKKVLESATWLEGDPYTGQK | 639 |
| Qy | 300 | YPSNGQGWGVTLDKSLNVAFAVNETSLSTNQKATYFTAQSGKPLKISLWSDAPASTS | 359 |
| Db | 640 | YTLEDQOQHGLVNVTKSWEI-----LKAINGTTLFIVDHWDKSYSDFF | 681 |
| Qy | 360 | A---SVTLVNDLVLITAPN-----GTKYGVN | 383 |
| Db | 682 | AEVLGVDVIRGLYARNSIPDIVEWHIKYVGD | 712 |

RESULT 5

| ID | EXPR | XANCP | STANDARD; | PRT; | 580 AA. |
|----|--|-------|-----------|------|---------|
| AC | P23314; | | | | |
| DT | 01-NOV-1991 (Rel. 20, Created) | | | | |
| DT | 01-NOV-1991 (Rel. 20, Last sequence update) | | | | |
| DT | 28-FEB-2003 (Rel. 41, Last annotation update) | | | | |
| DE | Extracellular protease precursor (EC 3.4.21.-). | | | | |
| GN | XCC0851. | | | | |
| OS | Xanthomonas campestris (pv. campestris). | | | | |
| OC | Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales; | | | | |
| OC | Xanthomonadaceae; Xanthomonas. | | | | |
| OX | NCBI_TaxID=340; | | | | |
| RP | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RX | MEDLINE=90251253; PubMed=2187155; | | | | |
| RA | Liu Y.-N., Tang J.-L., Clarke B.R., Dow J.M., Daniels M.J.; | | | | |
| RT | "A multipurpose broad host range cloning vector and its use to | | | | |
| RT | characterise an extracellular protease gene of Xanthomonas campestris | | | | |
| RT | pathovar campestris."; | | | | |
| RL | Mol. Gen. Genet. 220:433-440(1990). | | | | |
| RN | [2] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | STRAIN=ATCC 33913 / NCPPB 528; | | | | |
| RX | MEDLINE=22022145; PubMed=12024217; | | | | |
| RA | da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furian L.R., | | | | |
| RA | Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F., | | | | |
| RA | Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A., | | | | |
| RA | Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P., | | | | |
| RA | Ciccarelli R.B., Coutinho L.D., Cursino-Santos J.R., El-Dorri H., | | | | |
| RA | Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T., | | | | |
| RA | Forghieri E.F., Franco M.C., Greggio C.C., Gruber A., | | | | |
| RA | Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F., | | | | |
| RA | Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M., | | | | |
| RA | Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H., | | | | |
| RA | Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R., | | | | |
| RA | Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F., | | | | |
| RA | Spinoza L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D., | | | | |
| RA | Trindade dos Santos M., Truffi J., | | | | |
| RA | Setubal J.C., Kitajima J.P.; | | | | |
| RT | "Comparison of the genomes of two Xanthomonas pathogens with differing | | | | |
| RT | host specificities"; | | | | |
| RL | Nature 417:459-463(2002). | | | | |
| CC | -!- SUBCELLULAR LOCATION: Secreted. | | | | |
| CC | -!- SIMILARITY: Belongs to peptidase family S8. | | | | |
| CC | ----- | | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation | | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announcement | | | | |
| CC | or send an email to license@isb-sib.ch). | | | | |
| CC | ----- | | | | |
| DR | EMBL; X51635; CAA35962.1; -- | | | | |
| DR | EMBL; AF012184; XAM40166.1; -- | | | | |
| DR | PIR; S11890; S11890. | | | | |
| DR | HSPG; P00782; 2GST. | | | | |
| DR | MEROPS; S08.UFA; -- | | | | |

```

DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR InterPro; IPR009020; Protease_inhib.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydroxylase; Serine protease; Zymogen; Signal; Complete proteome.
FT SIGNAL 1 32
FT PROPEP 33 2136
FT CHAIN 2137 580
FT ACT_SITE 177 177
FT ACT_SITE 237 237
FT ACT_SITE 409 409
FT DISULFID 225 273
FT DISULFID 315 352
FT DISULFID 450 454
SQ SEQUENCE 580 AA; 57228 MW; 8C9A2CAE4E7F47CB CRC64;

Query Match 12.2%; Score 274; DB 1; Length 580;
Best Local Similarity 27.7%; Pred. No. 2.4e-10;
Matches 130; Conservative 53; Mismatches 157; Indels 130; Gaps 25;

QY 21 GGGQVAVADTGL-----DTGRDSSNVHEAFRGKITAIYALGRNNAND----- 64
DB 168 GSGTVAVDTGITSHADLANILAGYDFISDATTARDGNGRDSNADEGDWYAANEGCA 227

QY 65 -----PNGHGHVAG---SVLNGTSNKGMAPQANLVFQSVMSNGG----- 103
DB 228 GIPAASSSSWHGHVAGTAAVTNNTTGVAGTAYGAKVFPVRYLGRKCGGSLSDIADAIYMA 287

QY 104 ----LGLPLSNV--STLFQOAYSAGARIHTNSWGPVNGAYTTDSRNVDDYVRKNDMAVL 157
DB 288 SGGTVSGIFANPNAPVNIWS:GGGSCSTTQWNA-INGAVSRGT-----TVV 334

QY 158 FAAGNEGPGGTISAPGTAKNAITVQATEN--LRPSFGSYADNINHVAQFSRGTQDGR 215
DB 335 VAAGNDASNVSG-SLPANCANVIAVAATTSAGAKASYSNFGTGI----- 377

QY 216 IKPDWMAPGTFLSARSS--LAPDSSFANHDSKYAYMGTSMATPIVAGNVAQLREHFI 273
DB 378 ----DVSAPOSSILSTUNSTTTGSG-----ASVASYNGTSMASPHVAGVVALVQS--V 425

QY 274 KNRGITPK--PSLLK--AALIAGATDIGLYPSGNGQGRVTLDKSLNVAFVNETS---- 325
DB 426 APTALTFAAVETLLKNTARALPGAC-----SGCGGAGIVNADAATAVTA--INGSGCGG 477

QY 326 ---SLSTNQKATYSFTAQSGKPLKISLWSDAPASTASVTLV-----NDLDELVI---TA 374
DB 478 GGGNTLTNGTPVTGLGAATGAELNYTIT---VPAG-SGTLTVTTSGGSGDADLYVRAGSA 533

QY 375 PNGTKYVGNDFIAPYDNNMDGRNNVENVFINAPQSGTYTVEVQAYNVPOG 424
DB 534 PTDSATY---CHPYRS-----GNAETCTITAP-SGTYYVRUKAYSTSG 573

RESULT 6
SUBT_BACS9
ID SUBT_BACS9 STANDARD; PRT; 420 AA.
AC P28842;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Subtilisin precursor (EC 3.4.21.62).
GN SUB1.
OS Bacillus sp. (strain TA39).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=29336;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92256481; PubMed=1581352;

```

```

RA Narinx E., Davail S., Feller G., Gerday C.;
RT "Nucleotide and derived amino acid sequence of the subtilisin from
RT the antarctic psychrotroph Bacillus TA39.";
RL Biochim. Biophys. Acta 1131:111-113 (1992).
CC -I- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
CC it catalyzes the hydrolysis of proteins and peptide amides.
CC -I- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large uncharged residue
CC in P1. Hydrolyzes peptide amides.
CC -I- COFACTOR: Binds 1 calcium ion per subunit (Potential).
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- MISCELLANEOUS: Still active at temperatures close to 0 degrees
CC Celsius; it has a marked heat lability.
CC -I- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
CC sporulation, and many mutations which block sporulation at early
CC stages affect expression levels of subtilisin. However, subtilisin
CC is not necessary for normal sporulation.
CC -I- SIMILARITY: Belongs to peptidase family S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X62369; CAA44227.1; -.
DR EMBL; S23407; S23407.
DR HSP; Q99405; IMPT.
DR MEROPS; S08.00A; -.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydroxylase; Sporulation; Serine protease; Zymogen; Metal-binding;
KW Calcium-binding; Signal.
FT SIGNAL 1 26
FT PROPEP 27 111
FT CHAIN 112 420
FT ACT_SITE 145 145
FT ACT_SITE 182 182
FT ACT_SITE 360 360
FT METAL 115 115
FT METAL 154 154
SQ SEQUENCE 420 AA; 44086 MW; AE4F121BD32B26EC CRC64;

Query Match 12.1%; Score 272; DB 1; Length 420;
Best Local Similarity 30.4%; Pred. No. 2.1e-10;
Matches 96; Conservative 41; Mismatches 123; Indels 56; Gaps 15;

QY 21 GGGQVAVADTGLDTRNDSSMHEAFRGKITAI--VALGRT---NNANDPNGHGHVAGS 75
DB 136 GGGINIAVLDTGVTN-----HPDLNNVNEQCKDTVTGTTTNNSTCDRQGHGHVAGS 189

QY 76 VL---GNGTSKGMAPQANLVFQSVMSNGELGSLFNSVSTLFSQAYSAGARHTN--- 128
DB 190 ALADGGTNGVGVADPADLWAYKVLGDDG--SGYADDIAAIRHAGDQATALTQKVIN 247

QY 129 -SWGAPVNGAYTTDSRNVDDYVRKNDMAVLFAAGNEGPGGTISAPGTAKNAITVQATEN 187
DB 248 MSLGSGGESSLIITAVN---YSYNGKGLIIIAAAGNSGPGYQSGYPCALVNAVAALEN 304

QY 188 LRPSFGSYADNINHVAQFSRGT-KDG-----RIKPDWMAPGTFLSARSLAPSSFW 241
DB 305 -KVENGTY-----RVADFSSRGYSWTGDYVAIOKGDVEISAPGAAYST-----W 348

QY 242 ANHDSKYAYMGTSMATPIVAGNVAQLREHFIKNRGITPKPSLLKALIAAGATDIGLYP 301
DB 349 --FDGGVATISGTSMASPHAGLAAGLAAKIWAQYPSASNVDRGELQYRAY---ENDILSGY 403

```

```

QY 302 SG-----NQGWGRVLT 312
Db 404 AGYGDDFASGGGFAIV 419

RESULT 7
THES_BACSP
ID_THES_BACSP STANDARD; PRT; 401 AA.
AC 045670.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Thermophilic serine proteinase precursor (EC 3.4.21.-) (AK.1
DE protease).
OS Bacillus sp. (strain AK1).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95085262; PubMed=7993087;
RA MacIver B., McHale R.H., Saul D.J., Bergquist P.L.;
RT "Cloning and sequencing of a serine proteinase gene from a
RT thermophilic Bacillus species and its expression in Escherichia
RT coli.";
RL Appl. Environ. Microbiol. 60:3981-3988(1994).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.80 ANGSTROMS) OF 122-401.
RX MEDLINE=20057863; PubMed=10589904;
RA Smith C.A., Toogood H.S., Baker H.M., Daniel R.M., Baker E.N.;
RT "Calcium-mediated thermostability in the subtilisin superfamily: the
RT crystal structure of Bacillus Ak.1 protease at 1.8-A resolution.";
RL J. Mol. Biol. 294:1027-1040(1999).
CC -!- COFACTOR: Binds 3 calcium ions and 1 sodium ion per subunit.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: Has a pH optimum of 8.5, a temperature optimum of
CC 75 degrees Celsius.
CC -!- SIMILARITY: Belongs to peptidase family S8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L29506; AAA63688.1; -.
CC PIR; I39974; I39974.
CC PDB; 1DBI; 18-NOV-99.
CC
CC MEROPS; S08.009; -.
CC
CC InterPro; IPR000209; Peptidase_S8.
CC InterPro; IPR009020; Peptidase_Inhib.
CC Pfam; PF00082; Peptidase_S8; 1.
CC PRINTS; PR00723; SUBTILISIN.
CC PROSITE; PS00136; SUBTILASE ASP; 1.
CC PROSITE; PS00137; SUBTILASE HIS; 1.
CC PROSITE; PS00138; SUBTILASE_SER; 1.
CC
CC Hydrolase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
KW Signal; 3D-structure.
FT SIGNAL 1 24 POTENTIAL.
FT PROPEP 25 121
FT CHAIN 122 401
FT ACT_SITE 160 160
FT ACT_SITE 193 193
FT ACT_SITE 347 347
FT METAL 126 126
FT METAL 168 168
FT METAL 169 169
FT METAL 171 171
FT METAL 179 179
FT METAL 184 184
FT METAL 186 186
FT METAL 204 204

QY 204 204
FT METAL 207
FT METAL 209
FT METAL 211
FT METAL 297
FT METAL 300
FT METAL 323
FT METAL 323
FT DISULFID 258
FT TURN 128
FT TURN 128
FT HELIX 129
FT TURN 135
FT HELIX 137
FT TURN 140
FT TURN 141
FT TURN 145
FT TURN 152
FT TURN 152
FT STRAND 155
FT TURN 165
FT TURN 168
FT TURN 173
FT STRAND 174
FT TURN 180
FT STRAND 184
FT STRAND 184
FT HELIX 193
FT STRAND 219
FT TURN 228
FT HELIX 234
FT TURN 247
FT STRAND 251
FT HELIX 263
FT TURN 275
FT STRAND 278
FT STRAND 285
FT TURN 298
FT STRAND 301
FT TURN 309
FT STRAND 312
FT TURN 314
FT STRAND 316
FT TURN 320
FT STRAND 324
FT STRAND 331
FT TURN 336
FT STRAND 339
FT HELIX 346
FT TURN 363
FT HELIX 367
FT TURN 377
FT STRAND 379
FT TURN 382
FT STRAND 384
FT TURN 385
FT STRAND 388
FT STRAND 391
FT HELIX 395
FT TURN 400
SQ SEQUENCE 401 AA; 42835 MW; 1C736EF4A89F256F CRC64;

Query Match 11.8%; Score 264.5; DB 1; Length 401;
Best Local Similarity 32.6%; Pred. No. 5.9e-10;
Matches 86; Conservative 28; Mismatches 95; Indels 55; Gaps 11;

QY 11 DVAQSSGLYGGQVVAVADTGLDTRNDSSMHEAFRGKITAIYALGRNNDP---NG 67
Db 146 DVTKGS-----SQEIAVIDTGVD-----YTHPDLGKVIKGYDF--VDNDYDPMDLNN 192
QY 68 HGTHVAG---SVLNGCTSNKGAPQANLYFQSVMSNGGLGGLPSNVSTLFSQAYSAGAR 124
Db 193 HGTHVAGIAAAETNNATGAGMAPNTRILAVRALDRNG--SGTSLDIAAIYVADSAGAE 250
QY 125 IHTNSWGPVNGAYTTDSRVDYVKNDYAVLFAAGNEGPNGGTISAPCTAKNAITVGA 184
Db 251 VINLSLGC---DCHTTTLENVAVNYAMNKGSVVVAAGNNG--SSTTFEPASYENVIAGVA 305

CALCIUM 3.
CALCIUM 1.
CALCIUM 1 (VIA CARBONYL OXYGEN).
CALCIUM 1 (VIA CARBONYL OXYGEN).
SODIUM (VIA CARBONYL OXYGEN).
SODIUM (VIA CARBONYL OXYGEN).
SODIUM.

```

```

QY 185 TENLRPSGYSADININHVAFSGRGTGDKGR:KPDVMAFGTTLFSGARSLAPDSFWANH 244
Db 306 -----VDQYDLASFSNYGTW-----VDVAVGVIVSTITG----- 337
QY 245 DSKYAYMGTSMTATPVAGNVAQL 268
Db 338 -NRYAYMGTSMTASPRVAGLALL 360

RESULT 8
AQL1 THEAQ
ID AQL1 THEAQ STANDARD; PRT; 513 AA.
AC P08594;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aqualysin I precursor (EC 3.4.21.-).
GN PSTI.
OS Thermus aquaticus.
OC Bacteria; Deinococcus-Thermus; Thermococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=271;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 15-23.
RC STRAIN=Yt1;
RX MEDLINE=90216674; PubMed=2182621; Matsuzawa H., Ohta T.;
RA Terada I., Kwon S.-T., Miyata Y., Matsuzawa H., Taguchi H.,
RT "Unique precursor structure of an extracellular protease, aqualysin
RT I, with NH2- and COOH-terminal pro-sequences and its processing in
RT Escherichia coli."
RL J. Biol. Chem. 265:6576-6581(1990).
RN [2]
RP SEQUENCE OF 75-442 FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=Yt1;
RX MEDLINE=88225062; PubMed=3286255;
RA Kwon S.-T., Terada I., Matsuzawa H., Ohta T.;
RT "Nucleotide sequence of the gene for aqualysin I (a thermophilic
RT alkaline serine protease) of Thermus aquaticus YT-1 and
RT characteristics of the deduced primary structure of the enzyme."
RL Eur. J. Biochem. 173:491-497(1988).
RN [3]
RP SEQUENCE OF 128-170.
RX MEDLINE=88151937; PubMed=3162211;
RA Matsuzawa H., Tokugawa K., Hamaoki M., Mizoguchi M., Taguchi H.,
RA Terada I., Kwon S.-T., Ohta T.;
RT "Purification and characterization of aqualysin I (a thermophilic
RT alkaline serine protease) produced by Thermus aquaticus YT-1."
RL Eur. J. Biochem. 171:441-447(1988).
CC -1- FUNCTION: Aqualysin I is a thermophilic alkaline serine protease.
CC The optimal temperature for its caseinolytic activity is 80
CC degrees Celsius.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DEVELOPMENTAL STAGE: Secreted from the early stationary phase
CC until the time the cells cease to grow.
CC -1- PTM: THE N- AND C-TERMINAL PRO-SEQUENCES ARE REMOVED THROUGH THE
CC PROTEOLYTIC ACTIVITY OF AQUALYSIN I ITSELF. IN THAT ORDER, THE
CC C-TERMINAL PRO-SEQUENCE IS REQUIRED FOR TRANSLOCATION OF THE
CC PROTEASES ACROSS THE OUTER MEMBRANE.
CC -1- PTM: Two disulfide bonds are present.
CC -1- SIMILARITY: Belongs to peptidase family S8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90108; BAA14135.1; -
CC EMBL; X07734; CAA30559.1; -
CC EMBL; A35742; A35742.
CC HSP; P06873; 2PRK.

```

```

DR MEROPS; S08.051; -
DR InterPro; IPR000209; Peptidase S8.
DR InterPro; IPR009020; Protease Inhib.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
KW Hydrolase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 14
FT PROPEP 15 127
FT CHAIN 128 408 AQUALYSIN I.
FT PROPEP 409 513
FT ACT_SITE 166 166 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 197 197 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 349 349 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 513 AA; 53913 MW; DDFDFED4A50B785 CRC64;

Query Match 11.6%; Score 260.5; DB 1; Length 513;
Best Local Similarity 28.2%; Pred. No. 1.5e-09;
Matches 122; Conservative 43; Mismatches 157; Indels 111; Gaps 23;

QY 16 SYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITAIY-ALGRVTNANDPNHGHTHVAG 74
Db 152 TTTATGRGVNVYVDTGIRT-----THREFGGRARVGYDALG--GNGQDCNGHGHVAG 203
QY 75 SVLNGNTSNKGMAPOANLVPQSVYMDNSGGLGLPSNVSTLFSQAYSAGARIHTNSGAP- 133
Db 204 TI---GGVTGVAKAVNLVAVRVDLCNG--SGTSGV-----TAGVDWVTRHRRPA 250
QY 134 -----VNGAYTTDSRNVDDYVRKNDMAV---LFAAGNEGPNGGTISAPGTAKNAITVGA 184
Db 251 VANMSLGGGVST---ALDNAV-KNSIAAGVYVAAAGNDNANACNYS-PARVAEALTVGA 305
QY 185 T--ENLRPFSGYADNINHVAFSGRGTGDKGR:KPDVMAFGTTLFSGARSLAPDSFWA 242
Db 306 TTSSDAKASFSNYGSCV-----DLFAPGASIFSA-----WY 336
QY 243 NHDSKAYMGTSMTATPVAGNVAQLREHFINKRGITPKPSLLKAALIAAGT-----DIG 297
Db 337 TSDTATQTTLNGTSMTATPVAGVAALYLE---QNPST--PASVASAILNGATTGRLSGIG 391
QY 298 LGYP-----SGNQGWGRVTLDKSLNVAFVNETSLSLTNQKATYSFTAQSGKPLKISL 349
Db 392 SGSPNRLYLLSLSSGSGSTAPCTSCSYTGSLSGSDYNPQPNGTYYIS-----PAGTHR 446
QY 350 VMSDAPASTSASVTLVNDLVLITAPNGTKY--VGNDFTAPYDNNWDGRNNVENFINAP 407
Db 447 AMLRGPAGT-----DFDLVLRWDGSRWLTVGSS-TGP-----TSESLSYSG 488
QY 408 QSGTYTVEVQAYN 420
Db 489 TAGYLLMRIYAYS 501

RESULT 9
WPR1_BACSU STANDARD; PRT; 894 AA.
AC P54423; O06726;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cell wall-associated protease precursor (EC 3.4.21.-) [Contains: Cell
DE wall-associated polypeptides CWBP23 and CWBP52].
GN WPR1 OR BSU10770.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-54 AND 414-428.
RC STRAIN=168;
RX MEDLINE=97158234; PubMed=9004506;
RA Margot P., Karamata D.;

```


RA Remerowski M.L., Pepermans H.A.M., Hilbers C.W., van de Ven F.J.M.;
 RT "backbone dynamics of the 269-residue protease Savinase determined
 from 15N-NMR relaxation measurements.";
 RL Eur. J. Biochem. 235:629-640(1996).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (0.78 ANGSTROMS).
 RX MEDLINE=98426039; PubMed=9753430;
 RA Kuhn P., Knapp M., Soltis S.M., Ganshaw G., Thoenes M., Bott R.;
 RT "The 0.78-A structure of a serine protease: *Bacillus lentus*
 subtilisin.";
 RL Biochemistry 37:13446-13452(1998).
 CC -!- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
 it catalyzes the hydrolysis of proteins and peptide amides.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
 for peptide bonds, and a preference for a large uncharged residue
 in P1. Hydrolyzes peptide amides.
 CC -!- COFACTOR: Binds 2 calcium ions per subunit.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- BIOTECHNOLOGY: Used as a detergent protease. Sold under the name
 Savinase by Novozymes.
 CC -!- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
 sporulation, and many mutations which block sporulation at early
 stages affect expression levels of subtilisin. However, subtilisin
 is not necessary for normal sporulation.
 CC -!- SIMILARITY: Belongs to peptidase family S8.
 DR PDB; 1C9U; 06-OCT-99.
 DR PDB; 1C9N; 10-JAN-01.
 DR PDB; 1C9N; 10-JAN-01.
 DR PDB; 1GCI; 11-NOV-98.
 DR PDB; 1IAV; 11-JUL-01.
 DR PDB; 1JEA; 26-NOV-97.
 DR PDB; 1SVN; 14-OCT-96.
 DR MEROPS; S08.103; -.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Hydrolyase; Sporulation; Serine protease; Metal-binding;
 KW Calcium-binding; 3D-structure.
 FT ACT_SITE 32
 FT ACT_SITE 62
 FT ACT_SITE 215
 FT METAL 2
 FT METAL 40
 FT METAL 73
 FT METAL 75
 FT METAL 77
 FT METAL 79
 FT METAL 163
 FT METAL 165
 FT METAL 168
 FT STRAND 2
 FT STRAND 6
 FT HELIX 11
 FT TURN 12
 FT HELIX 13
 FT TURN 19
 FT TURN 20
 FT TURN 24
 FT TURN 25
 FT TURN 27
 FT TURN 32
 FT TURN 39
 FT STRAND 43
 FT TURN 51
 FT TURN 52
 FT HELIX 62
 FT STRAND 78
 FT TURN 84
 FT TURN 85
 FT STRAND 87
 FT STRAND 94
 FT STRAND 96
 FT STRAND 100
 FT HELIX 102
 FT TURN 114
 FT TURN 115
 FT STRAND 119

FT STRAND 126
 FT HELIX 131
 FT TURN 142
 FT STRAND 146
 FT STRAND 150
 FT STRAND 161
 FT TURN 162
 FT TURN 166
 FT STRAND 168
 FT STRAND 174
 FT TURN 176
 FT STRAND 180
 FT TURN 183
 FT TURN 188
 FT STRAND 190
 FT STRAND 199
 FT STRAND 203
 FT TURN 204
 FT STRAND 206
 FT TURN 207
 FT HELIX 214
 FT TURN 233
 FT HELIX 237
 FT TURN 246
 FT TURN 247
 FT STRAND 249
 FT HELIX 254
 FT TURN 257
 FT TURN 258
 FT STRAND 261
 FT HELIX 264
 FT TURN 268
 FT TURN 269
 SQ SEQUENCE 269 AA; 26698 MW; 4D89F8778999BF8D CRC64;
 Query Match 11.4%; Score 257; DB 1; Length 269;
 Best Local Similarity 30.3%; Pred. No. 1.1e-09;
 Matches 91; Conservative 38; Mismatches 117; Indels 54; Gaps 11;
 Qy 8 VKADVAQSSVGLYGQGVAVADTGLDTGENDSSMHEAPRGKITAIYALGRNTNANDPNG 67
 Db 11 VQAPAAHNR-GLTGGSGVKVAVLDTGIST-----HPDLNIRGASFFVGPGE-STDGNG 61
 Qy 68 HGHVAGSV--LNGTSTNKGMAPQANLVFQSVNDSNGGLGLPSNVSTLFSQAYSAGARI 125
 Db 62 HGHVAGTIAALNNSICVLGVAAPSALYAKVVLGASG--SGSVSSIAQGLEWAGNNGMHV 119
 Qy 126 HTNSWAGPVNGAYTTDSRNVDDIVRKNDMAVLPAAGNEPNGGTISAPGTAKNAITVGAT 185
 Db 120 ANLSLGSFSPSATLEQAVN---SATSRGVLVWAASGNSG--AGSISYPARYANAMAVGAT 174
 Qy 186 E--NLRFSGSYADNINNHVAQFSRGTGDKRIKPDVMAFGTIFLSARSLAPDSSFWAN 243
 Db 175 DQNNRASFSQYGAGL-----DIVAFGVNVQSTYPG----- 205
 Qy 244 HDSKIAYMGTSMATPIVAGNVAGLREHFKNGITPKPSLKAAALAGATDGLGYPG 303
 Db 206 --STYASLNGTSMATPHVAGAAALVKQKNPSWNVQIRNHLKNTATSLGNTL---YGS 260
 RESULT 11
 ID ELYA_BACAO STANDARD; PRT; 380 AA.
 AC P27693;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Alkaline protease precursor (EC 3.4.21.-).
 OS *Bacillus alcalophilus*.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
 OX NCBI_TaxID=1445;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PB92;
 RX MEDLINE=91282483; PubMed=2059048;
 RA van der Jaan J.C., Gerritse G., Mulleners L.J.M., van der Hoek R.A.,
 RA Quax W.J.;
 RT "Cloning, characterization, and multiple chromosomal integration of a
Bacillus alkaline protease gene."
 RL Appl. Environ. Microbiol. 57:901-909(1991).

RN X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
 RC STRAIN=PB92;
 RX MEDLINE=92390330; PubMed=1518788;
 RA van der Laan J.C., Teplyakov A.V., Kelders H., Kalk K.H., Misset O.,
 RA Mulleners L.J.M., Dijkstra B.W.;
 RT "Crystal structure of the high-alkaline serine protease PB92 from
 RT Bacillus alcalophilus.";
 RL Protein Eng. 5:405-411(1992).
 RN [3]
 RN X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
 RP MEDLINE=93078250; PubMed=1447775;
 RA Sobek H., Hecht H.-J., Lehle W., Schomburg D.;
 RT "X-ray structure determination and comparison of two crystal forms of
 RT a variant (Asn153Arg) of the alkaline protease from Bacillus
 RT alcalophilus refined at 1.85-A resolution.";
 RL J. Mol. Biol. 228:108-117(1992).
 RN [4]
 RN STRUCTURE BY NMR OF 112-380.
 RP STRAIN=PB92;
 RX MEDLINE=92727237; PubMed=9115441;
 RA Martin J.R., Mulder F.A., Karimi-Nejad Y., van der Zwan J.,
 RA Mariani M., Schipper D., Boelens R.;
 RT "The solution structure of serine protease PB92 from Bacillus
 RT alcalophilus presents a rigid fold with a flexible substrate-binding
 RT site.";
 RL Structure 5:521-532(1997).
 CC -!- COFACTOR: Binds 2 calcium ions per subunit.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to peptidase family S8.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M65086; AAA22212.1; -;
 CC EMBL; A13738; CAA01128.1; -;
 CC PIR; A49778; A49778.
 CC PDB; 1AH2; 15-APR-98.
 CC MEROPS; S08.038; -;
 CC InterPro; IPR000209; Peptidase S8.
 CC InterPro; IPR009020; Protease inhib.
 CC Pfam; PFO0082; Peptidase S8; 1.
 CC PRINTS; PR00723; SUBTILISIN.
 CC PROSITE; PS00136; SUBTILASE ASP; 1.
 CC PROSITE; PS00137; SUBTILASE HIS; 1.
 CC PROSITE; PS00138; SUBTILASE SER; 1.
 CC Hydrolase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
 KW Hydrolase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
 KW Signal; 3D-structure.
 FT SIGNAL 1 27 POTENTIAL.
 FT PROPEP 28 112
 FT CHAIN 113 380 ALKALINE PROTEASE.
 FT ACT_SITE 143 143 CHARGE RELAY SYSTEM.
 FT ACT_SITE 173 173 CHARGE RELAY SYSTEM.
 FT ACT_SITE 326 326 CHARGE RELAY SYSTEM.
 FT METAL 113 113 CALCIUM 1.
 FT METAL 151 151 CALCIUM 1.
 FT METAL 184 184 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 186 186 CALCIUM 1.
 FT METAL 188 188 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 190 190 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 274 274 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT METAL 276 276 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT METAL 279 279 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT METAL 117 120
 FT HELIX 121 123
 FT TURN 125 130
 FT HELIX 131 131
 FT TURN 135 136

FT STRAND 138 142
 FT TURN 150 151
 FT STRAND 156 157
 FT HELIX 173 182
 FT STRAND 198 201
 FT TURN 207 208
 FT HELIX 213 226
 FT STRAND 230 233
 FT HELIX 242 252
 FT TURN 253 255
 FT STRAND 257 258
 FT TURN 267 268
 FT TURN 272 275
 FT STRAND 279 279
 FT STRAND 282 285
 FT STRAND 291 291
 FT TURN 299 300
 FT STRAND 303 306
 FT STRAND 310 314
 FT TURN 315 317
 FT STRAND 318 322
 FT HELIX 325 342
 FT TURN 344 345
 FT HELIX 348 358
 FT HELIX 375 380
 SQ SEQUENCE 380 AA; 38853 MW; 539EA72771B6682C CRC64;
 Query Match 11.4%; Score 257; DB 1; Length 380;
 Best Local Similarity 30.3%; Pred. No. 1.7e-09;
 Matches 91; Conservative 38; Mismatches 117; Indels 54; Gaps 11;
 QY 8 VKADVAOSGLYGGQGVAVADTGLDTRNDSSMEAFRGKITAIYALGRTNNDPNP 67
 DB 122 VQAPAHNR-GLTSGGVKVAVLDTGIST-----HPDLNIRGGASFVPGEP-STQDNG 172
 QY 68 HGTHVAGSV--LNGTSTNKGWAPCANLVFQSVMDNSGGLGLPSNVSTLPSQAYSAGARI 125
 DB 173 HGTHVAGTIAALNNSIGVLGVAPNAELIYAVKVLGASG--SGSVSSIAQGLEWAGNNGMHV 230
 QY 126 HTNSGAPVNGAVYTTDSRNVDDYVRKNDMAVLFAAGNEGNGGTISAPGTAKNAITVGAT 185
 DB 231 ANLSLGSFSPSATLEQAVN--SATSRGVLVVAASGNSG--AGSISYPARYANAMAVGAT 285
 QY 186 E--NLRFSGSYADNHNHVAQFSRGTGDKRIKPDVMAPTFILSARSLAPDSSFWAN 243
 DB 286 DQNNRASFSQYGAGL-----DIVAPGVNVQSTYPG----- 316
 QY 244 HDSKYVMGTTSMATPIVAGNVAQLREHFINKRGITPKPSLLKALITAGATDIGLYPSG 303
 DB 317 --STYASLNGTSMATPHVAGAAALVKQKNFSWNVQIRNHLKNTATSLGNTL---YGS 371
 RESULT 12
 ID ELYA_BACCS STANDARD; PRT; 380 AA.
 AC P41362;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Alkaline protease precursor (EC 3.4.21.-).
 OS Bacillus clausii.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 CX NCBI TaxID=79880;
 RN [1]
 SEQUENCE FROM N.A.
 RP STRAIN=221 / ATCC 21522 / JCM 9139 / DSM 2512;
 RC MEDLINE=93043753; PubMed=1368952;
 RX Takami H., Kobayashi T., Kobayashi M., Yamamoto M., Nakamura S.,
 RA Aono R., Horikoshi K.;
 RT "Molecular cloning, nucleotide sequence, and expression of the
 RT structural gene for alkaline serine protease from alkaliphilic
 RT Bacillus sp. 221.";
 RL Biosci. Biotechnol. Biochem. 56:1455-1460(1992).

RN [2]
 RP SEQUENCE OF 112-129.
 RC STRAIN=221 / ATCC 25122 / JCM 9139 / DSM 2512;
 RA Horikoshi K.;
 RL (in) Horikoshi K. (eds.);
 RL Microorganisms in alkaline environments, pp.187-194, VCH,
 RL Weinheim (1991).
 CC -!- COFACTOR: Binds 2 calcium ions per subunit (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to peptidase family S8.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; S48754; AAC60420.1; -;
 DR EMBL; D13157; BAA02442.1; -;
 DR EMBL; A26817; BAA01836.1; -;
 DR EMBL; A22550; CAA01611.1; -;
 DR HSSP; P29600; IGC1.
 DR MEROPS; S08.103; -;
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR009020; Protease_inhib.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE SER; 1.
 DR Hydrolase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
 KW SIGNAL.
 FT SIGNAL 1 27 POTENTIAL.
 FT PROPEP 28 111
 FT CHAIN 112 380
 FT ACT_SITE 143 143 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 173 173 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 326 326 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT METAL 113 113 CALCIUM 1 (BY SIMILARITY).
 FT METAL 151 151 CALCIUM 1 (BY SIMILARITY).
 FT METAL 184 184 CALCIUM 1 (VIA CARBONYL OXYGEN)
 FT METAL 186 186 CALCIUM 1 (BY SIMILARITY).
 FT METAL 188 188 CALCIUM 1 (VIA CARBONYL OXYGEN)
 FT METAL 190 190 CALCIUM 1 (VIA CARBONYL OXYGEN)
 FT METAL 274 274 CALCIUM 2 (VIA CARBONYL OXYGEN)
 FT METAL 276 276 CALCIUM 2 (VIA CARBONYL OXYGEN)
 FT METAL 279 279 CALCIUM 2 (VIA CARBONYL OXYGEN)
 FT METAL 279 279 CALCIUM 2 (VIA CARBONYL OXYGEN)
 FT METAL 279 279 CALCIUM 2 (VIA CARBONYL OXYGEN)
 FT METAL 279 279 CALCIUM 2 (VIA CARBONYL OXYGEN)
 SQ SEQUENCE 380 AA; 38826 MW; 5F73ABC68D5B6831 CRC64;
 Query Match 11.4%; Score 257; DB 1; Length 380;
 Best Local Similarity 30.3%; Pred. No. 1.7e-09;
 Matches 91; Conservative 38; Mismatches 117; Indels 54; Gaps 11;
 QY 8 VKADVAQSSGLYGGQGVAVADTGLDGTGRDMSMEAFKGTATYALGRTNANDPNG 67
 DB 122 VQFAAENR-GLTSGVKKVAVLTGTST-----HPDLNIRGGSFVPGEF-STODGNG 172
 QY 68 HGTAVAGSV--LNGTSKNGMAPANLVFQSVMDNSGGLGLPSNVSTLFSQAYSAGARI 125
 DB 173 HGTAVAGTIAALNNSIGVLGVAEALYAVKVLGASG--SGSVSSIAQGLEWAGNGMHV 230
 QY 126 HTNSWGPVNGAYTTDSRNVDDYVYKNDMAVLPAAGNEGNGGTISAPGTAKVAITVGAT 185
 DB 231 ANLSLGSPSPSATLEQAVN---SATSRGVLVVAASGNSG--AGSISYPARYANAMAVGAT 285

QY 186 E--NLRPSEGSYADNINHVAFSPSRGPTKDGRIKPDVMAPTFILSARSSLAPDSSFWAN 243
 DB 286 DQNNNRASFQYAGL-----DIVAPGVNQSTYFG----- 316
 QY 244 HDSKYAYMGTSMAPIVAGNVAQVREHFKNRGITPKPSLLKAALIAAGATDGLGYPG 303
 DB 317 --STVASLNGTSMATPHVAGAAALVKQKPNVSNVQIRNHLKNTATSLGSTNL---YSGG 371
 RESULT 13
 PRTM_BACSP STANDARD; PRT; 269 AA.
 AC Q99405; 1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE M-protease (EC 3.4.21.-)
 OS Bacillus sp. (strain KSM-K16).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1409;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RA Yamane T., Kani T., Hatanaka T., Suzuki A., Ashida T.,
 RA Kobayashi T., Ito S., Yamashita O.;
 RT "Structure of a new alkaline serine protease (M-protease) from
 RT Bacillus sp. KSM-K16".
 RL Acta Crystallogr. D 51:199-206 (1995).
 RN [2]
 RP SEQUENCE OF 1-23, AND CHARACTERIZATION.
 RX MEDLINE=9535832; PubMed=7632397;
 RA Kobayashi T., Hakamada Y., Adachi S., Hitomi J., Yoshimatsu T.,
 RA Koike K., Kawai S., Ito S.;
 RT "Purification and properties of an alkaline protease from
 RT alkalophilic Bacillus sp. KSM-K16".
 RL Appl. Microbiol. Biotechnol. 43:473-481 (1995).
 CC -!- COFACTOR: Binds 2 calcium ions per subunit.
 CC -!- SUBUNIT: Monomer.
 CC -!- SIMILARITY: Belongs to peptidase family S8.
 DR MEROPS; S08.010; -;
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE SER; 1.
 DR Hydrolase; Serine protease; Metal-binding; Calcium-binding;
 KW 3D-structure.
 FT ACT_SITE 32 32 CHARGE RELAY SYSTEM.
 FT ACT_SITE 62 62 CHARGE RELAY SYSTEM.
 FT ACT_SITE 215 215 CHARGE RELAY SYSTEM.
 FT METAL 2 2 CALCIUM 1.
 FT METAL 40 40 CALCIUM 1.
 FT METAL 73 73 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 75 75 CALCIUM 1.
 FT METAL 77 77 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 79 79 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 163 163 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT METAL 165 165 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT METAL 168 168 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT HELIX 6 10
 FT TURN 11 12
 FT TURN 13 18
 FT TURN 19 20
 FT TURN 24 25
 FT STRAND 27 32
 FT TURN 39 40
 FT STRAND 43 48
 FT TURN 51 52
 FT TURN 62 71
 FT TURN 84 85
 FT TURN 87 92
 FT TURN 96 97


```

FT HELIX 102 114
FT TURN 115 116
FT STRAND 119 122
FT STRAND 126 126
FT HELIX 131 142
FT TURN 143 144
FT STRAND 146 150
FT STRAND 161 161
FT TURN 162 164
FT TURN 166 167
FT STRAND 174 174
FT TURN 176 177
FT STRAND 180 180
FT TURN 182 183
FT TURN 188 189
FT STRAND 192 195
FT STRAND 199 203
FT TURN 204 206
FT STRAND 207 211
FT HELIX 214 231
FT TURN 233 234
FT HELIX 237 246
FT TURN 247 247
FT STRAND 249 249
FT HELIX 254 257
FT TURN 258 259
FT STRAND 261 261
FT HELIX 264 267
FT TURN 268 269
SQ SEQUENCE 269 AA; 26723 MW; 7A03C86D534A1D07 CRC64;

Query Match 11.3%; Score 255; DB 1; Length 269;
Best Local Similarity 30.4%; Pred. No. 1.4e-09;
Matches 96; Conservative 32; Mismatches 102; Indels 86; Gaps 14;

QY 8 VKADVAQSSYGLGQGVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTNANDPNG 67
Db 11 VQAPRAHNR-GLTSGVKVAVLDTGIST-----HPDLNRGGASFPVGEPT-STDQNG 61

QY 68 HGHVAGSV--LNGTSNKGMAPQANLVFQSVMDNSGGLGLPSNVSTLFSQAVSAGARI 125
Db 62 HGHVAGTITAAALNSIGVLGAPSAELXAVKVLGASG--SGSVSIAQGLEWAGNGMHV 119

QY 126 HTNSGAPVNGYATYDTSRNVDDYVRKNDMAVLFAAGNEGPGGTTISAPGTAKNAITVGAT 185
Db 120 ANLSLGSPPSPTLQAVN---SATSRGVLVVAASGNSG--AGSISYPARVANAVAGAT 174

QY 186 E--NLRPFGSVADNINHVAFSPSRGPTKDGRIKPDVMAFGTFLSARSSLAPDSSFWAN 243
Db 175 DONNRASPSQYAGL-----DIVAPGVNVQSTYFG-----205

QY 244 HDSKAYMGSTMATPIVAGNVA-----QLREHFTKRGITPKPSLLKAALIA 291
Db 206 --STVASLNGTSMATPHVAGVAAVLKQKNPSNVQIRNH-LKN-----246

QY 292 GATDIGLG----YPSG 303
Db 247 --TATGLGNTLYGSG 260

RESULT 14
PROA_VIBAL STANDARD; PRT; 534 AA.
AC F16588;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alkaline serine exoprotease A precursor (EC 3.4.21.-).
GN PROA.
OS Vibrio alginolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=663;

[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=89326126; PubMed=2546861;
RX Deane S.M., Robb F.T., Robb S.M., Woods D.R.;
RT "Nucleotide sequence of the Vibrio alginolyticus calcium-dependent,
RL detergent-resistant alkaline serine exoprotease A.";
RL Gene 76:281-288(1989).
CC -|- SIMILARITY: Belongs to peptidase family S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M25499; AAA27550.1; -.
DR PIR; J50173; J50173.
DR HSP; Q99405; IMPT.
DR MEROPS; S08.050; -.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR InterPro; IPR009020; Protease_inhib.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
KW Hydrolase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 21
FT PROPEP 22 141
FT CHAIN 142 534
FT ACT_SITE 180 180
FT ACT_SITE 213 213
FT ACT_SITE 363 363
SQ SEQUENCE 534 AA; 55930 MW; 84E96D9C649D4226 CRC64;

Query Match 11.2%; Score 252.5; DB 1; Length 534;
Best Local Similarity 26.1%; Pred. No. 5e-09;
Matches 122; Conservative 61; Mismatches 151; Indels 133; Gaps 24;

QY 7 IVKADVAQSS--YGL-----YGGQGVAVADTGLDTRNDSSMHEAF 46
Db 137 IVSADANQNAIWGLDRIDORNLPDNNYSANFDGTGTAYVIDGV-----NNAHVEF 190

QY 47 RGKITAIY-ALGRTNANDPENGHTHVAGSVLNGTSNKGMAPQANLVFQSVN--DSNGG 103
Db 191 GGRSVSGYDFVNDADASDCNGHGHVAGTI---GGSLYGVAKVNVNLVGRVLSGSGS 247

QY 104 LGGLPSNVSTLFSQAVSAGARIHTNSGAPVNGAYTDSRNVDDYVR---KNDMAVLFAA 160
Db 248 TSQVIAGVDWVAANA--SGPSVANMGLGGQSV-----LDSAVQSAVQSGVSFLAA 298

QY 161 GNEGPGGTTISAPGTAKNAITVGAT--ENLRPFGSVADNINHVAFSPSRGPTKDGRIKP 218
Db 299 GNSNADACNYS-PARVATGTVGTTSTTDARSFSNWGSCV-----338

QY 219 DVMAPGTFLSARSSLAPDSSFWANHDSKAYMGSTMATPIVAGNVAOLREHFIKNGRI 278
Db 339 DVFAPGSQIKSA-----W-YDGGYKXTISGTSMATPHVAG-VAAL--YLQENSSV 383

QY 279 TPKPSLLKAALITAGA-----TDIGLGFSGNGGWRVTLDKSLNVA 319
Db 384 S--PSQVEALIVSRASGTGKVTDRGSKVLLXSLTDADCGQCGGPD---PTPDPEGKLT 438

QY 320 FVNETSLL--STNQKATYFTQSGPKLISLWSDAPASTASVTLVNDLIDLVITAPNG 377
Db 439 SGVPVSLGSGSGQVAYYYVDVEAGQRLTVQMYGS-----GDADLYLRF--G 484

QY 378 TKVVGNDFTAPYDNNWDGR-----NNVENVFINAPQSGTYTVEVQAYN 420

```

Db 485 AK-----PTLNAWDCRPFYKGNCTCTVSATQSGRYHYMIQGS 523

RESULT 15

SUBT_BACLI STANDARD; PRT; 379 AA.

ID P00780; AC

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Subtilisin Carlsberg precursor (EC 3.4.21.62).

GN APR

OS Bacillus licheniformis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1402;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NCIB 6816;

RX MEDLINE=6093688; PubMed=3001653;

RA Jacobs M., Eliasson M., Uhlen M., Flock J.-I.,

RT "Cloning, sequencing and expression of subtilisin Carlsberg from

RT Bacillus licheniformis.";

RL Nucleic Acids Res. 13:8913-8926(1985).

RN [2]

RP SEQUENCE OF 106-379.

RX MEDLINE=68234702; PubMed=4967581;

RA Smith E.L., Delange R.J., Evans W.H., Landon M., Markland F.S.;

RT "Subtilisin Carlsberg. V. The complete sequence; comparison with

RT subtilisin BPN'; evolutionary relationships.";

RL J. Biol. Chem. 243:2184-2191(1968).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF MUTANT WITH

RP SELENOCYSTEINE-325

RX MEDLINE=93291170; PubMed=8512925;

RA Syed R., Wu Z.F., Hogle J.M., Hilvert D.;

RT "Crystal structure of selenosubtilisin at 2.0-A resolution.";

RL Biochemistry 32:6157-6164(1993).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 106-379.

RX MEDLINE=98087517; PubMed=9425066;

RA Stoll V.S., Eger B.T., Hynes R.C., Martichonok V., Jones J.B.,

RA Pai E.F.;

RT "Differences in binding modes of enantiomers of L-acetamidoboronic

RT acid based protease inhibitors: crystal structures of gamma-

RT chymotrypsin and subtilisin Carlsberg complexes.";

RL Biochemistry 37:451-462(1998).

CC -I- FUNCTION: Subtilisin is an extracellular alkaline serine protease,

CC it catalyzes the hydrolysis of proteins and peptide amides.

CC -I- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity

CC for peptide bonds, and a preference for a large uncharged residue

CC in P1. Hydrolyzes peptide amides.

CC -I- COFACTOR: Binds 2 calcium ions per subunit.

CC -I- SUBCELLULAR LOCATION: Secreted.

CC -I- BIOTECHNOLOGY: Used as a detergent protease. Sold under the name

CC Alcalase by Novozymes.

CC -I- MISCELLANEOUS: Secretion of subtilisin is associated with onset of

CC sporulation, and many mutations which block sporulation at early

CC stages affect expression levels of subtilisin. However, subtilisin

CC is not necessary for normal sporulation.

CC -I- SIMILARITY: Belongs to peptidase family S8.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: X03341; CAB56500.1; ..

CC PIR: A24111; SUBSCU.

CC PDB: 1AF4; 16-JUN-97.

CC PDB: 1AV7; 01-APR-98.

CC -----

DR PDB; 1AV7; 25-MAR-98.

DR PDB; 1BE6; 14-OCT-98.

DR PDB; 1BB3; 13-JAN-99.

DR PDB; 1BFK; 18-NOV-98.

DR PDB; 1BTU; 18-NOV-98.

DR PDB; 1C31; 04-AUG-99.

DR PDB; 1CSE; 15-OCT-89.

DR PDB; 1SBC; 15-JAN-95.

DR PDB; 1SCA; 31-JAN-94.

DR PDB; 1SCB; 31-JAN-94.

DR PDB; 1SCD; 31-JAN-94.

DR PDB; 1SCN; 31-AUG-94.

DR PDB; 1SEL; 31-OCT-93.

DR PDB; 1VSB; 18-MAR-98.

DR PDB; 2SEC; 15-JAN-95.

DR PDB; 3VSB; 25-MAR-98.

DR MEROPS; S08.001; ..

DR InterPro; IPR000209; Peptidase_S8.

DR InterPro; IPR009020; Protease_inhib.

DR Pfam; PF00082; Peptidase_S8; 1.

DR PRINTS; PR00723; SUBTILISIN.

DR PROSITE; PS00136; SUBTILASE_ASP; 1.

DR PROSITE; PS00137; SUBTILASE_HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.

DR KW Hydrolase; Sporulation; Serine protease; Zymogen; Metal-binding;

KW Calcium-binding; Signal; 3D-structure.

FT SIGNAL 1 29

FT PROPEP 30 105

FT CHAIN 106 379

FT ACT_SITE 137 137

FT ACT_SITE 168 168

FT ACT_SITE 325 325

FT METAL 107 107

FT METAL 146 146

FT METAL 179 179

FT METAL 181 181

FT METAL 185 185

FT METAL 273 273

FT METAL 275 275

FT METAL 278 278

FT CONFLICT 207 207

FT CONFLICT 233 233

FT CONFLICT 262 265

FT CONFLICT 316 316

FT TURN 111 111

FT HELIX 112 116

FT TURN 117 117

FT TURN 118 124

FT TURN 125 125

FT TURN 129 130

FT STRAND 132 137

FT TURN 142 143

FT TURN 145 146

FT STRAND 149 154

FT TURN 157 158

FT TURN 161 162

FT HELIX 168 177

FT TURN 190 191

FT STRAND 193 198

FT TURN 202 203

FT STRAND 205 205

FT HELIX 208 220

FT TURN 221 222

FT STRAND 225 228

FT STRAND 230 232

FT HELIX 237 249

FT TURN 250 250

FT STRAND 252 256

FT TURN 264 265

FT STRAND 271 271

FT TURN 272 274

FT TURN 276 277

FT STRAND 279 284

```
FT TURN 286 287
FT STRAND 290 290
FT TURN 292 293
FT TURN 298 299
FT STRAND 302 305
FT STRAND 309 313
FT TURN 314 316
FT STRAND 317 321
FT HELIX 324 341
FT TURN 343 344
FT HELIX 347 356
FT TURN 357 357
FT STRAND 359 359
FT HELIX 364 367
FT TURN 368 369
FT STRAND 371 371
FT HELIX 374 377
FT TURN 378 379
SQ SEQUENCE 379 AA; 38908 MW; F19A6DC5761FB504 CRC64;

Query Match 11.1%; Score 250.5; DB 1; Length 379;
Best Local Similarity 30.2%; Pred. No. 4.3e-09;
Matches 92; Conservative 42; Mismatches 112; Indels 59; Gaps 12;

QY 7 IVKADVAOSSYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITAIYALGRTNNANDPN 66
DB 115 LIKADKVQQAQ-CFKGANVKVAVLDIGIQASHEDLNVVG-----ASFVAGEAYN-TDCN 166

QY 67 GHGTHVAGSV--LNGTSNKGMAPQANLVFSVMDNSGGLGLPSNVSTLFSQAYSAGAR 124
DB 167 GHGTHVAGTVAALDNTTGLGVAPSVSLYAVKVLNLSG--SGTYSGIVSGIEWATTNGMD 224

QY 125 IHTNSGAPVNGAYTDSRNVDDYVRKNDMAVLFAAGNEGNG--GTISAPGTAKNAITY 182
DB 225 VINMSLGGP---SGSTAMKQAVDNAYARGVVVAAAGNSGSGNTNTTIGYPAKYDSVIIV 281

QY 183 GATENLRPSFGSYADNINHVAQFSSRGPTKGRIXPDYMAPGTFTLSARSSLAPDSSFWA 242
DB 282 GA-----VDSNSNRASFSSVG-----AELEVMAPGAGVYSTYPT----- 315

QY 243 NHDSKYAYMGTSMATPIVAGNVAQLREHFIGNRITPKPSL-----LKAALIAGNTDIGL 298
DB 316 ---STVATLNGTSMASPHVAGAAALI-----LSKHPNLSASQVRNRUSSTATYLG 363

QY 299 GYPSG 303
DB 364 SFYVG 368
```

Search completed: March 31, 2004, 16:05:32
Job time : 8.51154 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 16:09:09 ; Search time 33.1412 Seconds
(without alignments)
3418.697 Million cell updates/sec

Title: US-09-985-689A-7
Perfect score: 2252
Sequence: 1 NDVARGIVKADVAQNFGLY.....EVQYNVFVSPQTFSLAIVH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 261661801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/prodata/1/pubaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/1/pubaa/US07_PUBCOMB.pep.*
3: /cgn2_6/prodata/1/pubaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/1/pubaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/1/pubaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/1/pubaa/US07_PUBCOMB.pep.*
7: /cgn2_6/prodata/1/pubaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/1/pubaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/1/pubaa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodata/1/pubaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/1/pubaa/US09_PUBCOMB.pep.*
12: /cgn2_6/prodata/1/pubaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/1/pubaa/US10_PUBCOMB.pep.*
14: /cgn2_6/prodata/1/pubaa/US10_PUBCOMB.pep.*
15: /cgn2_6/prodata/1/pubaa/US10_PUBCOMB.pep.*
16: /cgn2_6/prodata/1/pubaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/1/pubaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|---------------------|
| 1 | 2252 | 100.0 | 433 | 10 | US-09-985-689A-7 |
| 2 | 2125.5 | 94.4 | 434 | 10 | US-09-985-689A-1 |
| 3 | 2125.5 | 94.4 | 434 | 15 | US-10-385-662-2 |
| 4 | 2118.5 | 94.1 | 434 | 10 | US-09-985-689A-2 |
| 5 | 2116.5 | 94.0 | 434 | 10 | US-09-985-689A-5 |
| 6 | 2044 | 90.8 | 433 | 10 | US-09-985-689A-5 |
| 7 | 2040 | 90.6 | 433 | 10 | US-09-985-689A-3 |
| 8 | 2033 | 90.3 | 433 | 10 | US-09-985-689A-4 |
| 9 | 1649 | 73.2 | 345 | 14 | US-10-336-324-10 |
| 10 | 1649 | 73.2 | 345 | 14 | US-10-403-105-13 |
| 11 | 432.5 | 19.2 | 659 | 13 | US-10-090-624-12 |
| 12 | 396 | 17.6 | 412 | 13 | US-10-090-624-1 |
| 13 | 396 | 17.6 | 522 | 13 | US-10-090-624-1 |
| 14 | 396 | 17.6 | 654 | 13 | US-10-090-624-16 |
| 15 | 355 | 15.8 | 1208 | 14 | US-10-156-761-13251 |

Sequence 39, Appl
Sequence 10856, A
Sequence 4, Appl
Sequence 6, Appl
Sequence 114, Appl
Sequence 4, Appl
Sequence 55, Appl
Sequence 2, Appl
Sequence 59, Appl
Sequence 12934, A
Sequence 3, Appl
Sequence 255271
Sequence 2, Appl
Sequence 60, Appl
Sequence 56, Appl
Sequence 46, Appl
Sequence 48, Appl
Sequence 50, Appl
Sequence 52, Appl
Sequence 54, Appl
Sequence 44, Appl
Sequence 58, Appl
Sequence 10, Appl
Sequence 16, Appl
Sequence 1, Appl
Sequence 8, Appl
Sequence 49, Appl
Sequence 1, Appl

US-10-112-488-39
US-10-156-761-10856
US-10-314-657-4
US-10-090-624-6
US-10-084-846A-114
US-10-084-846A-4
US-09-927-827-55
US-09-966-921A-2
US-09-927-827-59
US-10-156-761-12934
US-10-344-231-3
US-10-424-599-255271
US-09-813-408-2
US-09-927-827-60
US-10-243-549-56
US-10-243-549-46
US-10-243-549-48
US-10-243-549-50
US-10-243-549-52
US-10-243-549-54
US-10-243-549-60
US-10-243-549-44
US-10-243-549-58
US-08-322-678-10
US-09-837-235-16
US-09-860-854B-6
US-09-975-139-1
US-09-976-414-8
US-09-736-116-49
US-10-075-907-1

ALIGNMENTS

RESULT 1

US-09-985-689A-7
Sequence 7, Application US/09985689A
Publication No. US20030022351A1
GENERAL INFORMATION: YUJI
APPLICANT: HATADA, AKINORI
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SATO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NORIYUKI
APPLICANT: OKUDA, MATSUYOSHI
APPLICANT: SAEKI, KATSUHI
TITLE OF INVENTION: Alkaline proteases
FILE REFERENCE: 215483USO
CURRENT APPLICATION NUMBER: US/09/985,689A
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: JP P2000-355166
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: JP P2001-114048
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 433
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689A-7

Query Match 100.0%; Score 2252; DB 10; Length 433;
Best Local Similarity 100.0%; Pred. No. 2.4e-189;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNFGLYQGQIVAVDTGLTGRNDSMHEAFRGKITALVALGRTN 60

Db 1 NDVARGIVKADVAQNFGLYQGQIVAVDTGLTGRNDSMHEAFRGKITALVALGRTN 60

QY 61 NANDPENGHTGVAGSVLGNATNKGMPOANLVFQISINDSGGLGLPANLQTLFSQAYSA 120

Db 61 NNDPKNHGHVAGSVLGNATKGMAPQANLVFQSDMSGGGLGGLPANLQTLFQAYS 120
Qy 121 GARIHNSGAPVNGAYTTDSRVDDYVRKNDMTILFAAGNEPGSGTISAPGTAKNAI 180
Db 121 GARIHNSGAPVNGAYTTDSRVDDYVRKNDMTILFAAGNEPGSGTISAPGTAKNAI 180
Qy 181 VGATENLRPSFGSYADNINHVAQFSSRGPTDRGRIPKPDVMAQGTIILSARSLAPDSSF 240
Db 181 VGATENLRPSFGSYADNINHVAQFSSRGPTDRGRIPKPDVMAQGTIILSARSLAPDSSF 240
Qy 241 ANHDSKYAYMGGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAAALTAGAADVGLGF 300
Db 241 ANHDSKYAYMGGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAAALTAGAADVGLGF 300
Qy 301 NGNGQGRVTLDSKLVAFNETSPSTOKATYSFTAQAGKPKISLVWSDAPGSTTAS 360
Db 301 NGNGQGRVTLDSKLVAFNETSPSTOKATYSFTAQAGKPKISLVWSDAPGSTTAS 360
Qy 361 LTLVNDLDELITAPNGTKYVGNDFTPAYDNNMDGRNNVENFINAPQSGTYTVEVQAYN 420
Db 361 LTLVNDLDELITAPNGTKYVGNDFTPAYDNNMDGRNNVENFINAPQSGTYTVEVQAYN 420
Qy 421 PVSPQTFSLAIVH 433
Db 421 PVSPQTFSLAIVH 433

RESULT 2

US-09-985-689A-1
; Sequence 1, Application US/09985689A
; Publication No. US20030022351A1

; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI

; APPLICANT: OGAWA, AKINORI

; APPLICANT: KAGEYAMA, YASUSHI

; APPLICANT: SATO, TSUYOSHI

; APPLICANT: ARAKI, HIROYUKI

; APPLICANT: SUMITOMO, NOBUYUKI

; APPLICANT: OKUDA, MITSUYOSHI

; APPLICANT: SAEKI, KATSUHIISA

; TITLE OF INVENTION: Alkaline proteases

; FILE REFERENCE: 215483USO

; CURRENT APPLICATION NUMBER: US/09/985,689A

; CURRENT FILING DATE: 2002-07-01

; PRIOR APPLICATION NUMBER: JP P2000-355166

; PRIOR FILING DATE: 2000-11-22

; PRIOR APPLICATION NUMBER: JP P2001-114048

; PRIOR FILING DATE: 2001-04-12

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 434

; TYPE: PRT

; ORGANISM: Bacillus sp.

US-09-985-689A-1

Query Match 94.4%; Score 2125.5; DB 10; Length 434;
Best Local Similarity 93.5%; Pred. No. 3.3e-178;
Matches 406; Conservative 19; Mismatches 8; Indels 1; Gaps 1;

Qy 1 NDVARGIVKADVQANNFGLYGQGOIVAVADTGLDGTGRNDSMHEAFRGKITALYALGRTN 60
Db 1 NDVARGIVKADVQANNFGLYGQGOIVAVADTGLDGTGRNDSMHEAFRGKITALYALGRTN 60

Qy 61 NNDPKNHGHVAGSVLGN-ATNKGMAPQANLVFQSDMSGGGLGGLPANLQTLFQAYS 119
Db 61 NNDPKNHGHVAGSVLGN-ATNKGMAPQANLVFQSDMSGGGLGGLPANLQTLFQAYS 120

Qy 120 AGARIHNSWGAAPVNGAYTTDSRVDDYVRKNDMTILFAAGNEPGSGTISAPGTAKNAI 179
Db 120 AGARIHNSWGAAPVNGAYTTDSRVDDYVRKNDMTILFAAGNEPGSGTISAPGTAKNAI 180

Qy 180 TVGATENLRPSFGSYADNINHVAQFSSRGPTDRGRIPKPDVMAQGTIILSARSLAPDSSF 239
Db 180 TVGATENLRPSFGSYADNINHVAQFSSRGPTDRGRIPKPDVMAQGTIILSARSLAPDSSF 240
Qy 240 WANHDSKYAYMGGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAAALTAGAADVGLGF 299
Db 240 WANHDSKYAYMGGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAAALTAGAADVGLGF 300
Qy 300 PNGNGQGRVTLDSKLVAFNETSPSTOKATYSFTAQAGKPKISLVWSDAPGSTTAS 359
Db 300 PNGNGQGRVTLDSKLVAFNETSPSTOKATYSFTAQAGKPKISLVWSDAPGSTTAS 360
Qy 360 SLTLVNDLDELITAPNGTKYVGNDFTPAYDNNMDGRNNVENFINAPQSGTYTVEVQAYN 419
Db 360 SLTLVNDLDELITAPNGTKYVGNDFTPAYDNNMDGRNNVENFINAPQSGTYTVEVQAYN 420
Qy 420 PVSPQTFSLAIVH 433
Db 420 PVSPQTFSLAIVH 434

RESULT 3

US-10-385-662-2

; Sequence 2, Application US/10385662

; Publication No. US20040002432A1

; GENERAL INFORMATION:

; APPLICANT: OKUDA, MITSUYOSHI

; APPLICANT: SATO, TSUYOSHI

; APPLICANT: SAITO, KAZUHIRO

; APPLICANT: SUMITOMO, NOBUYUKI

; APPLICANT: IZAWA, YOSHIFUMI

; APPLICANT: SAEKI, KATSUHIISA

; APPLICANT: KOBAYASHI, TOHRO

; APPLICANT: NOMURA, MASAFUMI

; TITLE OF INVENTION: Alkaline protease

; FILE REFERENCE: 234938USO

; CURRENT APPLICATION NUMBER: US/10/385,662

; CURRENT FILING DATE: 2003-03-12

; PRIOR APPLICATION NUMBER: JP 2002-081428

; PRIOR FILING DATE: 2002-03-22

; PRIOR APPLICATION NUMBER: JP 2002-165987

; PRIOR FILING DATE: 2002-06-06

; PRIOR APPLICATION NUMBER: JP 2002-304230

; PRIOR FILING DATE: 2002-10-18

; PRIOR APPLICATION NUMBER: JP 2002-304231

; PRIOR FILING DATE: 2002-10-18

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 434

; TYPE: PRT

; ORGANISM: Bacillus sp. KSM-KP43

US-10-385-662-2

Query Match 94.4%; Score 2125.5; DB 15; Length 434;
Best Local Similarity 93.5%; Pred. No. 3.3e-178;
Matches 406; Conservative 19; Mismatches 8; Indels 1; Gaps 1;

Qy 1 NDVARGIVKADVQANNFGLYGQGOIVAVADTGLDGTGRNDSMHEAFRGKITALYALGRTN 60
Db 1 NDVARGIVKADVQANNFGLYGQGOIVAVADTGLDGTGRNDSMHEAFRGKITALYALGRTN 60

Qy 61 NNDPKNHGHVAGSVLGN-ATNKGMAPQANLVFQSDMSGGGLGGLPANLQTLFQAYS 119
Db 61 NNDPKNHGHVAGSVLGN-ATNKGMAPQANLVFQSDMSGGGLGGLPANLQTLFQAYS 120

Qy 120 AGARIHNSWGAAPVNGAYTTDSRVDDYVRKNDMTILFAAGNEPGSGTISAPGTAKNAI 179
Db 120 AGARIHNSWGAAPVNGAYTTDSRVDDYVRKNDMTILFAAGNEPGSGTISAPGTAKNAI 180

Qy 180 TVGATENLRPSFGSYADNINHVAQFSSRGPTDRGRIPKPDVMAQGTIILSARSLAPDSSF 239
Db 180 TVGATENLRPSFGSYADNINHVAQFSSRGPTDRGRIPKPDVMAQGTIILSARSLAPDSSF 240

QY 240 WANHDSKYAYMGSTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAAGADVGLGF 299
 DB 241 WANHDSKYAYMGSTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAAGADVGLGF 300
 QY 300 PNGNQGWGRVTLDKSLNVAFNVTSPSTLSQKATYSFTAQAGKPLKISLVNSDAPGSTTA 359
 DB 301 PNGNQGWGRVTLDKSLNVAFNVTSPSTLSQKATYSFTAQAGKPLKISLVNSDAPGSTTA 360
 QY 360 SLTLVNDLDLVTAPNGTKYVGNDFTPAYDNNWGRNNVENVFINAPQSGTYTVEVQAYN 419
 DB 361 SVTLVNDLDLVTAPNGTKYVGNDFTPAYDNNWGRNNVENVFINAPQSGTYTVEVQAYN 420
 QY 420 VPVSPQTFSLAIVH 433
 DB 421 VPVGPQTFSLAIVN 434

RESULT 4

US-09-985-689A-2

; Sequence 2, Application US/09985689A

; Publication No. US20030022351A1

; GENERAL INFORMATION:

; APPLICANT: HATADA, YUJI

; APPLICANT: OGAWA, AKINORI

; APPLICANT: KAGEYAMA, YASUSHI

; APPLICANT: SATO, TSUYOSHI

; APPLICANT: ARAKI, HIROYUKI

; APPLICANT: SUMITOMO, NOBUYUKI

; APPLICANT: OKUDA, MITSUYOSHI

; APPLICANT: SAEKI, KATSUHIISA

; TITLE OF INVENTION: Alkaline proteases

; FILE REFERENCE: 215483US0

; CURRENT APPLICATION NUMBER: US/09/985,689A

; CURRENT FILING DATE: 2002-07-01

; PRIOR APPLICATION NUMBER: JP P2000-355166

; PRIOR FILING DATE: 2000-11-22

; PRIOR APPLICATION NUMBER: JP P2001-114048

; PRIOR FILING DATE: 2001-04-12

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 434

; TYPE: PRT

; ORGANISM: Bacillus sp.

US-09-985-689A-2

Query Match 94.1%; Score 2118.5; DB 10; Length 434;
 Best Local Similarity 93.3%; Pred. No. 1.3e-177;
 Matches 405; Conservative 18; Mismatches 10; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNPFGLYGQGI VAVADTGLDGRNDSMHEAFRGKITAYALGRTN 60
 DB 1 NDVARGIVKADVAQNPFGLYGQGI VAVADTGLDGRNDSMHEAFRGKITAYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGN-ATNKGMAPOANLVFQSDIMSDGGGLGGLPANLQTLFSQAYS 119
 DB 61 NANDPNHGHTHVAGSVLGNATNKGMAPOANLVFQSDIMSDGGGLGGLPSNVSTLFSQAYS 120
 QY 120 AGARIHTNSWGA PVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179
 DB 121 AGARIHTNSWGA PVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 180
 QY 180 TVGATENLRPSFGSYADNTHVAQFSRGPTRDGRIPKPDVMA PGTYILSARSLAPDSSF 239
 DB 181 TVGATENLRPSFGSYADNTHVAQFSRGPTRDGRIPKPDVMA PGTYILSARSLAPDSSF 240
 QY 240 WANHDSKYAYMGSTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAAGADVGLGF 299
 DB 241 WANHDSKYAYMGSTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAAGADVGLGF 300
 QY 300 PNGNQGWGRVTLDKSLNVAFNVTSPSTLSQKATYSFTAQAGKPLKISLVNSDAPGSTTA 359

DB 301 PNGNQGWGRVTLDKSLNVAFNVTSPSTLSQKATYSFTAQAGKPLKISLVNSDAPGSTTA 360
 QY 360 SLTLVNDLDLVTAPNGTKYVGNDFTPAYDNNWGRNNVENVFINAPQSGTYTVEVQAYN 419
 DB 361 SVTLVNDLDLVTAPNGTKYVGNDFTPAYDNNWGRNNVENVFINAPQSGTYTVEVQAYN 420
 QY 420 VPVSPQTFSLAIVH 433
 DB 421 VPVGPQTFSLAIVN 434

RESULT 5

US-09-985-689A-6

; Sequence 6, Application US/09985689A

; Publication No. US20030022351A1

; GENERAL INFORMATION:

; APPLICANT: HATADA, YUJI

; APPLICANT: OGAWA, AKINORI

; APPLICANT: KAGEYAMA, YASUSHI

; APPLICANT: SATO, TSUYOSHI

; APPLICANT: ARAKI, HIROYUKI

; APPLICANT: SUMITOMO, NOBUYUKI

; APPLICANT: OKUDA, MITSUYOSHI

; APPLICANT: SAEKI, KATSUHIISA

; TITLE OF INVENTION: Alkaline proteases

; FILE REFERENCE: 215483US0

; CURRENT APPLICATION NUMBER: US/09/985,689A

; CURRENT FILING DATE: 2002-07-01

; PRIOR APPLICATION NUMBER: JP P2000-355166

; PRIOR FILING DATE: 2000-11-22

; PRIOR APPLICATION NUMBER: JP P2001-114048

; PRIOR FILING DATE: 2001-04-12

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 434

; TYPE: PRT

; ORGANISM: Bacillus sp.

US-09-985-689A-6

Query Match 94.0%; Score 2116.5; DB 10; Length 434;
 Best Local Similarity 92.3%; Pred. No. 2e-177;
 Matches 400; Conservative 21; Mismatches 12; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNPFGLYGQGI VAVADTGLDGRNDSMHEAFRGKITAYALGRTN 60
 DB 1 NDVARGIVKADVAQNPFGLYGQGI VAVADTGLDGRNDSMHEAFRGKITAYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGNAT-NKGMAPOANLVFQSDIMSDGGGLGGLPANLQTLFSQAYS 119
 DB 61 NANDPNHGHTHVAGSVLGNATNKGMAPOANLVFQSDIMSDGGGLGGLPSNVSTLFSQAYS 120
 QY 120 AGARIHTNSWGA PVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179
 DB 121 AGARIHTNSWGA PVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 180
 QY 180 TVGATENLRPSFGSYADNTHVAQFSRGPTRDGRIPKPDVMA PGTYILSARSLAPDSSF 239
 DB 181 TVGATENLRPSFGSYADNTHVAQFSRGPTRDGRIPKPDVMA PGTYILSARSLAPDSSF 240
 QY 240 WANHDSKYAYMGSTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAAGADVGLGF 299
 DB 241 WANHDSKYAYMGSTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAAGADVGLGF 300
 QY 300 PNGNQGWGRVTLDKSLNVAFNVTSPSTLSQKATYSFTAQAGKPLKISLVNSDAPGSTTA 359
 DB 301 PNGNQGWGRVTLDKSLNVAFNVTSPSTLSQKATYSFTAQAGKPLKISLVNSDAPGSTTA 360
 QY 360 SLTLVNDLDLVTAPNGTKYVGNDFTPAYDNNWGRNNVENVFINAPQSGTYTVEVQAYN 419
 DB 361 SVTLVNDLDLVTAPNGTKYVGNDFTPAYDNNWGRNNVENVFINAPQSGTYTVEVQAYN 420
 QY 420 VPVSPQTFSLAIVH 433


```
/ APPLICANT: SUMITOMO, NOBUYUKI
/ APPLICANT: OKUDA, MITSUYOSHI
/ APPLICANT: SAEKI, KATSUHISA
/ TITLE OF INVENTION: Alkaline proteases
/ FILE REFERENCE: 215483USO
/ CURRENT APPLICATION NUMBER: US/09/985,689A
/ PRIOR FILING DATE: 2002-07-01
/ PRIOR APPLICATION NUMBER: JP P2000-355166
/ PRIOR FILING DATE: 2000-11-22
/ PRIOR APPLICATION NUMBER: JP P2001-114048
/ PRIOR FILING DATE: 2001-04-12
/ NUMBER OF SEQ ID NOS: 7
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 4
/ TYPE: PRT
/ ORGANISM: Bacillus sp.
US-09-985-689A-4

Query Match
Best Local Similarity 90.3%; Score 2033; DB 10; Length 433;
Matches 384; Conservative 25; Mismatches 24; Indels 0; Gaps 0;

Qy 1 NDVARGIVKADVAQNNGFLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
Db 1 NDVARGIVKADVAQNNGFLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
Qy 61 NANDPNHGHTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 120
Db 61 NANDPNHGHTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 120
Qy 121 GARIHTNSWGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
Db 121 GARIHTNSWGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPGSGTISAPGTAKNAIT 180
Qy 181 VGATENLRPSFGSYADNINHVAQFSRGTRDGR:KPDVMAPTGVILSARSSLAPDSFW 240
Db 181 VGATENLRPSFGSIADNHNHIAQFSRGATRDGR:KPDVTPAGTIFLSARSSLAPDSFW 240
Qy 241 ANHDSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAAIAGADVGLGFP 300
Db 241 ANYSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAAIAGATDVGLGYP 300
Qy 301 NGNQGWGRVTLDKSLNVAFVNSTPLSTSQKATYSTAQAGKPLKISLWSDAPGSTTAS 360
Db 301 NGDQGWGRVTLKSLNVAVNEATATLQKATYSTFQAQAGKPLKISLWTDAPGSTTAS 360
Qy 361 LTLVNDLDLVIITAPNGKTVGVNDFTAPYDNNWGDGNNVFNAPQSGTIVVEQAYNV 420
Db 361 YTLVNDLDLVIITAPNGKTVGVNDFTAPYDNNWGDGNNVFNAPQSGTIVVEQAYNV 420
Qy 421 PVSPQTFSLAIVH 433
Db 421 PSGPQRFSLAIVH 433

RESULT 9
US-10-336-324-10
/ Sequence 10, Application US/10336324
/ Publication No. US20030176304A1
/ GENERAL INFORMATION:
/ APPLICANT: Hansen, Peter
/ APPLICANT: Bauditz, Peter
/ APPLICANT: Mikkelsen, Frank
/ APPLICANT: Andersen, Kim
/ TITLE OF INVENTION: Protease Variants and Compositions
/ FILE REFERENCE: 5349,204-US
/ CURRENT APPLICATION NUMBER: US/10/336,324
/ CURRENT FILING DATE: 2003-01-03
/ PRIOR APPLICATION NUMBER: US/09/512,251A
/ PRIOR FILING DATE: 2000-02-24
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn version 3.1

Query Match
Best Local Similarity 73.2%; Score 1649; DB 14; Length 345;
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NDVARGIVKADVAQNNGFLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
Db 29 NDVARGIVKADVAQNNGFLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 88
Qy 61 NANDPNHGHTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 120
Db 89 NANDPNHGHTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 148
Qy 121 GARIHTNSWGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180

Query Match
Best Local Similarity 100.0%; Pred. No. 1.9e-136;
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NDVARGIVKADVAQNNGFLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
Db 29 NDVARGIVKADVAQNNGFLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 88
Qy 61 NANDPNHGHTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 120
Db 89 NANDPNHGHTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 148
Qy 121 GARIHTNSWGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180

Query Match
Best Local Similarity 73.2%; Score 1649; DB 14; Length 345;
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NDVARGIVKADVAQNNGFLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
Db 29 NDVARGIVKADVAQNNGFLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 88
Qy 61 NANDPNHGHTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 120
Db 89 NANDPNHGHTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 148
Qy 121 GARIHTNSWGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180

Query Match
Best Local Similarity 100.0%; Pred. No. 1.9e-136;
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NDVARGIVKADVAQNNGFLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
Db 29 NDVARGIVKADVAQNNGFLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 88
Qy 61 NANDPNHGHTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 120
Db 89 NANDPNHGHTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 148
Qy 121 GARIHTNSWGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
```


Db 149 GARIHTNSWGPVNGAYTTDSNRVDDYVRKNDMTILFAAGNEBPGSGTISAPGTAKNAIT 208
 Qy 181 VGATENLRPSFGSYADNINHAQFSSRGPTDRGRKPKDVMAPGTIYLSARSLAPDSSFW 240
 Db 209 VGATENLRPSFGSYADNINHAQFSSRGPTDRGRKPKDVMAPGTIYLSARSLAPDSSFW 268
 Qy 241 ANHDSKYAYMGTSMTATPIVAGNVAQLREHFVNKRGVTPKPSLLKAALIAAGAADVGLGFP 300
 Db 269 ANHDSKYAYMGTSMTATPIVAGNVAQLREHFVNKRGVTPKPSLLKAALIAAGAADVGLGFP 328
 Qy 301 NGNOGWGRVTLKSLNV 317
 Db 329 NGNOGWGRVTLKSLNV 345

RESULT 11
 US-10-090-624-12
 ; Sequence 12, Application US/10090624
 ; Publication No. US20020132335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TAKAKURA, Hikaru
 ; APPLICANT: MORISHITA, Mio
 ; APPLICANT: SHIMOJO, Tomoko
 ; APPLICANT: ASADA, Kiyozo
 ; APPLICANT: KATO, Ikunoshin
 ; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
 ; FILE REFERENCE: TAKAKURA=6
 ; CURRENT APPLICATION NUMBER: US/10/090,624
 ; CURRENT FILING DATE: 2002-03-06
 ; PRIOR APPLICATION NUMBER: 09/445,472
 ; PRIOR FILING DATE: 1999-12-06
 ; PRIOR APPLICATION NUMBER: 151969/1997
 ; PRIOR FILING DATE: 1997-06-10
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 12
 ; LENGTH: 659
 ; TYPE: PRT
 ; ORGANISM: Thermococcus celer
 US-10-090-624-12

Query Match 19.2%; Score 432.5; DB 13; Length 659;
 Best Local Similarity 30.7%; Pred. No. 4.6e-29;
 Matches 134; Conservative 62; Mismatches 156; Indels 85; Gaps 16;

Qy 8 VKADVAQNFLGYGQGIIVADTGLDGTGRNDSMHEAFRGKITALLY-ALGRNTNANDEN 66
 Db 145 IGADTVNWSLGYDGSVVVAIVDTGIDAN-----HPDLKGRVIGYDAVNGRSTPYDDQ 198
 Qy 67 GHGTHVAGSVLG-----NATNKGMAPOANLVFOSIM--DSGGGLGGLPANLQTLFSAQYSA 120
 Db 199 GHGTHVAGSVLG-----NATNKGMAPOANLVFOSIM--DSGGGLGGLPANLQTLFSAQYSA 120
 Qy 121 GARIHTNSWGPVNGAYTTDSNRVDDYVRKNDMTILFAAGNEBPGSGTISAPGTAKNAIT 180
 Db 259 GIRVINLSGSSQSSDGTSLQAVNNAWDAGLVVVAAGNSGPNKYTIGSPAAASKVITV 318
 Qy 181 VGATENLRPSFGSYADNINHAQFSSRGPTDRGRKPKDVMAPGTIYLSARSLAPDSSFW 240
 Db 319 VGA-----VDSNDNIAFSRSGPTADGRKPKDVMAPGTIYLSARSLAPDSSFW 364
 Qy 241 ANHDSKYAYMGTSMTATPIVAGNVAQLREHFVNKRGVTPK--PSLLKAALIAAG----- 292
 Db 365 TPINDYTKASGTSMA:PHVSGVGAALLQAH-----PSWTPDKVKYKTALETADIVAP 416
 Qy 293 ---ADVGLGFPNGQGRVTLKSL---NVAFVNETSPILTSQKATYSFTAQAGKPLKI 346
 Db 417 KEIADIAYGA-----GRVNVYKAIKYDDYAKLTFTGSAVADKGSATHTF----- 459
 Qy 347 SLVNSDAPGSTTASLTLY-----NDLDLVTAPNGTKYVGNDFTPADYDNNWDRNNVENV 401
 Db 460 -----DVSGATFVTATLYWDPTGSSDIDLVDLPNGNS-VDSYTYAY-----GFEKY 505

Qy 402 FINAQSGTYTVEVOAY 418
 Db 506 GYNNPTAGTWTKVVS 522

RESULT 12
 US-10-090-624-1
 ; Sequence 1, Application US/10090624
 ; Publication No. US20020132335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TAKAKURA, Hikaru
 ; APPLICANT: MORISHITA, Mio
 ; APPLICANT: SHIMOJO, Tomoko
 ; APPLICANT: ASADA, Kiyozo
 ; APPLICANT: KATO, Ikunoshin
 ; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
 ; FILE REFERENCE: TAKAKURA=6
 ; CURRENT APPLICATION NUMBER: US/10/090,624
 ; CURRENT FILING DATE: 2002-03-06
 ; PRIOR APPLICATION NUMBER: 09/445,472
 ; PRIOR FILING DATE: 1999-12-06
 ; PRIOR APPLICATION NUMBER: 151969/1997
 ; PRIOR FILING DATE: 1997-06-10
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 412
 ; TYPE: PRT
 ; ORGANISM: Pyrococcus furiosus
 US-10-090-624-1

Query Match 17.6%; Score 396; DB 13; Length 412;
 Best Local Similarity 29.5%; Pred. No. 3.7e-26;
 Matches 127; Conservative 56; Mismatches 163; Indels 84; Gaps 15;

Qy 16 NFGLYGQGIIVADTGLDGTGRNDSMHEAFRGKITALLYALGRNTNAN-----DENGHG 69
 Db 20 NLGYDGSGITIGIDTGID-----ASHPDLQKGV-----IGHVDFVNGRSVPYDDHGHG 68
 Qy 70 THVAGSVLG-----NATNKGMAPOANLVFOSIM--DSGGGLGGLPANLQTLFSAQYSA 122
 Db 69 THVASIAAGTGAASNGYKQVAPGAKLAGIKVLGADSGSISITIIKGVEWAVDNKDKYGI 128
 Qy 123 RIHTNSWGPVNGAYTTDSNRVDDYVRKNDMTILFAAGNEBPGSGTISAPGTAKNAITVG 182
 Db 129 KVINLSLGSQSSDGTSLQAVNNAWDAGLVVVAAGNSGPNKYTIGSPAAASKVITV 188
 Qy 183 ATENLRPSFGSYADNINHAQFSSRGPTDRGRKPKDVMAPGTIYLSARSLAPDSSFWAN 242
 Db 189 A-----VDKYDVIITFSRSGPTADGRKPKDVMAPGTIYLSARSLAPDSSFWAN 234
 Qy 243 HDSKYAYMGTSMTATPIVAGNVAQLREHFVNKRGVTPK--PSLLKAALIAAG----- 292
 Db 235 INDYTAAPGTSMTATPIVAGNVAQLREHFVNKRGVTPK--PSLLKAALIAAG----- 287
 Qy 293 ADVGLGFPNGQGRVTLKSL---NVAFVNETSPILTSQKATYSFTAQAGKPLKISLV 349
 Db 288 ADIAYGA-----GRVNVYKAIKYDDYAKLTFTGSAVADKGSATHTF----- 340
 Qy 350 WSDAPGSTTASLTLYWDPTGSSDIDLVDLPNGNS-VDSYTYAY-----GFEKY 409
 Db 341 WDNAN-----SDLDLYLVDPNQNG-VDSYTYAY-----GFEKY 381
 Qy 410 TYTVEVOAYN 419
 Db 382 TWTKVVSYS 391

RESULT 13
 US-10-090-624-4
 ; Sequence 4, Application US/10090624
 ; Publication No. US20020132335A1
 ; GENERAL INFORMATION:

APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOJO, Tomoko
APPLICANT: ASADA, Kiyocho
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/10/090,624
CURRENT FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: 09/445,472
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patent in version 3.0
SEQ ID NO 4
LENGTH: 522
TYPE: PRT
ORGANISM: Pyrococcus furiosus
FEATURE:
NAME/KEY: misc feature
LOCATION: (428)..(428)
OTHER INFORMATION: Xaa at position 428 is Gly or Val.
US-10-090-624-4

Query Match 17.6%; Score 396; DB 13; Length 522;
Best Local Similarity 29.5%; Pred. No. 5.3e-26;
Matches 127; Conservative 56; Mismatches 163; Indels 84; Gaps 15;
QY 16 NFGLVGQGIIVAVDTGLDTGRNDSMHEAFRGKITALLYALGRITNNAN-----DPNGHG 69
DB 20 NLGYDGSGITIGIITGID-----ASHPDLQGV-----IGWDFVNGRSYPYDDHGHG 68
QY 70 THVAGSVLG-----NATNKGMAPOANLVFQSIM--DSGGGLGLPANLQTLFSQAYSAGA 122
DB 69 THVASTAAGTGAASNGKYKGMAPGAKLAGIKVLGADGSGSISTIIKGVWAVDNKDKYGI 128
QY 123 RIHTNSWGPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAITVG 182
DB 129 KVINLSLGSQSSDGTALSOAVNAADAGLVVVVAAGNSGPNKYTIGSPAAASKVITVG 188
QY 183 ATENLRPSGVSADNINHVAQFSSRGPTDGRKPDVMAPGTYILSARSSSLAPDSSFWAN 242
DB 189 A-----VDKYDVITSSRGPTADGRKPEVAPGNWIIAARAS---GTSMGQP 234
QY 243 HPSKYAYMGTSMTATPIVAGNVAQREHFVKNRGVTPK--PSLLKAALIAGA----- 292
DB 235 INDYTAAPGTSMTATPIVAGNVAQREHFVKNRGVTPK--PSLLKAALIAGA----- 287
QY 293 ADVGLGFPNGQGWGRVTLDKSL---NVAFVNETSPLTSQKATYSFTAQAGKPLKISLV 349
DB 288 ADIAYGA-----GRVNAKAINVDYAKLVFTGYVANKGSGTHQFVSGASFTVATLY 340
QY 350 WSDAPGSTTASLTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGRNNVNFVINAPOSQ 409
DB 341 WDNAN-----SDLDLYLDPNGNQ--VDYSYTAYY-----GFEKVGYNPTD 381
QY 410 TYTVEVOAYN 419
DB 382 TWIKVVSYS 391

RESULT 14
US-10-090-624-16
Sequence 16, Application US/10090624
Publication No. US2002032335A1
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOJO, Tomoko
APPLICANT: ASADA, Kiyocho
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE

FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/10/090,624
CURRENT FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: 09/445,472
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patent in version 3.0
SEQ ID NO 16
LENGTH: 654
TYPE: PRT
ORGANISM: Pyrococcus furiosus
US-10-090-624-16

Query Match 17.6%; Score 396; DB 13; Length 654;
Best Local Similarity 29.5%; Pred. No. 7.3e-26;
Matches 127; Conservative 56; Mismatches 163; Indels 84; Gaps 15;
QY 16 NFGLVGQGIIVAVDTGLDTGRNDSMHEAFRGKITALLYALGRITNNAN-----DPNGHG 69
DB 152 NLGYDGSGITIGIITGID-----ASHPDLQGV-----IGWDFVNGRSYPYDDHGHG 200
QY 70 THVAGSVLG-----NATNKGMAPOANLVFQSIM--DSGGGLGLPANLQTLFSQAYSAGA 122
DB 201 THVASTAAGTGAASNGKYKGMAPGAKLAGIKVLGADGSGSISTIIKGVWAVDNKDKYGI 260
QY 123 RIHTNSWGPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAITVG 182
DB 261 KVINLSLGSQSSDGTALSOAVNAADAGLVVVVAAGNSGPNKYTIGSPAAASKVITVG 320
QY 183 ATENLRPSGVSADNINHVAQFSSRGPTDGRKPDVMAPGTYILSARSSSLAPDSSFWAN 242
DB 321 A-----VDKYDVITSSRGPTADGRKPEVAPGNWIIAARAS---GTSMGQP 366
QY 243 HPSKYAYMGTSMTATPIVAGNVAQREHFVKNRGVTPK--PSLLKAALIAGA----- 292
DB 367 INDYTAAPGTSMTATPIVAGNVAQREHFVKNRGVTPK--PSLLKAALIAGA----- 419
QY 293 ADVGLGFPNGQGWGRVTLDKSL---NVAFVNETSPLTSQKATYSFTAQAGKPLKISLV 349
DB 420 ADIAYGA-----GRVNAKAINVDYAKLVFTGYVANKGSGTHQFVSGASFTVATLY 472
QY 350 WSDAPGSTTASLTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGRNNVNFVINAPOSQ 409
DB 473 WDNAN-----SDLDLYLDPNGNQ--VDYSYTAYY-----GFEKVGYNPTD 513
QY 410 TYTVEVOAYN 419
DB 514 TWIKVVSYS 523

RESULT 15
US-10-156-761-13251
Sequence 13251, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109

```
; SEQ ID NO 13251
; LENGTH: 1208
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13251

Query Match          15.8%; Score 355; DB 14; Length 1208;
Best Local Similarity 31.7%; Pred. No. 7,2e-22;
Matches 126; Conservative 51; Mismatches 151; Indels 70; Gaps 17;

QY 8 VKADVAQNN-----FGLYGQGIIVAVADTGLDTRNDSSWHEAFRGKITALYALG 57
Db 186 VEADMAESNAQIGTRAAMDAGLTGDSVTIVAVLDTGVDT-----THPDLAGRVSRSKSF 239
QY 58 RTNNANDPNHGHTHVAGSVLG-----NATNKGMAPQANLVFQSIW-DSGGGLGLPANLQ 111
Db 240 DGEVADRNGHGTHTVTSIVGGSGAASDGTGCVAFGATLAVGKVLSDQAG-----SESQ 294
QY 112 TLFSQAYS-----GARIHTNSWAPVNGAYTTD--SRNVDDYVRKNDMTILFAAGNEGPGS 166
Db 295 ILAGMEWAARDVRARIVSMGLGS-TEASDGTDPMAEAVDTLSEETGALFVVAAGNTGAPS 353
QY 167 GTISAGTKNAITVCATENLRPFSGSYADNINHVAQFSSRGPTR-DGRIKPDVWAPGTY 225
Db 354 -SIGSPGAADSALTUGA-----VDSDRARYFTSAGPRHGDNALKPEDLAAPGYD 401
QY 226 ILSARSSLAPDSFWMANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLK 285
Db 402 IRAARSQLAPGTGY-----YTSMSGTSMATPHVAGVAALLAECHPDWTGARLKDALMS 454
QY 286 AA--LIAGRADVCLGFPNGQNGH-RVTLDKSLNVAF-----VNETSPLSTSOKA 332
Db 455 TSQOLDASVYQLGAGRVSPDVGARVATATGSADLGFHFWPHDADRPTKTVTYSNSDT 514
QY 333 TYSFT-AQAGKPLKISLVMSDA---PGSTTASLTLVND 366
Db 515 TVELSLAVRGAPAGVATLADTALTVPAGHTAATTVTGD 552
```

Search completed: March 31, 2004, 16:34:03
Job time : 34.1412 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:38 ; Search time 11.3246 Seconds
(without alignments)
3677.911 Million cell updates/sec

Title: US-09-985-689A-7
Perfect score: 2252
Sequence: 1 NDVARGIVKADVAQNNGLY.....EVQANVFPVQTFSLAIVH 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|--------|---------------------|
| 1 | 499 | 22.2 | 1743 | T18279 | multidrug resistan |
| 2 | 494.5 | 22.0 | 1905 | T18267 | multidrug resistan |
| 3 | 334.5 | 14.9 | 444 | B83891 | intracellular alka |
| 4 | 316.5 | 14.1 | 442 | A69587 | intracellular alka |
| 5 | 309 | 13.7 | 806 | A41341 | microbial serine p |
| 6 | 304 | 13.5 | 1388 | T28159 | pyrolysin (EC 3.4. |
| 7 | 283.5 | 12.6 | 799 | G83753 | subtilisin-type pr |
| 8 | 279.5 | 12.4 | 580 | S11890 | serine proteinase |
| 9 | 277.5 | 12.3 | 1345 | T29090 | surface layer-asso |
| 10 | 266 | 11.8 | 420 | S23407 | subtilisin (EC 3.4 |
| 11 | 265 | 11.8 | 419 | S25835 | subtilisin (EC 3.4 |
| 12 | 262.5 | 11.7 | 715 | JC4908 | alkaline serine pr |
| 13 | 261.5 | 11.6 | 627 | D75393 | serine proteinase, |
| 14 | 258 | 11.5 | 757 | C84120 | subtilisin-type pr |
| 15 | 258 | 11.5 | 1331 | A72647 | probable surface l |
| 16 | 254 | 11.3 | 534 | J50173 | alkaline proteinase |
| 17 | 249 | 11.1 | 401 | I39974 | serine proteinase |
| 18 | 245.5 | 10.9 | 321 | S27501 | alkaline proteinase |
| 19 | 245.5 | 10.9 | 380 | A49778 | high-alkaline seri |
| 20 | 245.5 | 10.9 | 513 | A35742 | aqualysin (EC 3.4. |
| 21 | 243.5 | 10.8 | 894 | F69730 | cell wall-associat |
| 22 | 240.5 | 10.7 | 382 | I39780 | subtilisin (EC 3.4 |
| 23 | 238 | 10.6 | 525 | G84406 | halolysin (importe |
| 24 | 237.5 | 10.5 | 519 | S71451 | halolysin R4 (EC 3 |
| 25 | 234.5 | 10.4 | 910 | C69456 | subtilisin sendai |
| 26 | 231.5 | 10.3 | 769 | D86335 | T20H2.6 protein - |
| 27 | 230 | 10.2 | 326 | C41335 | microbial serine p |
| 28 | 230 | 10.2 | 1167 | A35066 | streptococcal C5a |
| 29 | 229.5 | 10.2 | 319 | I39866 | microbial serine p |

| | | | | | | |
|----|-------|------|------|---|--------|---------------------|
| 30 | 229 | 10.2 | 402 | 1 | JU0332 | alkaline proteinase |
| 31 | 229 | 10.2 | 488 | 2 | A11930 | proteinase [import |
| 32 | 228 | 10.1 | 1052 | 2 | T17093 | intraluminal subti |
| 33 | 227.5 | 10.1 | 378 | 2 | A33973 | high-alkaline seri |
| 34 | 226 | 10.0 | 379 | 1 | SUBSCL | subtilisin (EC 3.4 |
| 35 | 226 | 10.0 | 1433 | 1 | A36734 | bacillopeptidase F |
| 36 | 225.5 | 10.0 | 706 | 2 | T01351 | subtilisin-like pr |
| 37 | 224.5 | 10.0 | 1374 | 2 | D72593 | hypothetical prote |
| 38 | 224 | 9.9 | 401 | 2 | A57690 | aerolysin precurs |
| 39 | 223.5 | 9.9 | 333 | 2 | I39867 | microbial serine p |
| 40 | 223 | 9.9 | 530 | 2 | A42605 | halolysin (EC 3.4. |
| 41 | 222 | 9.9 | 382 | 1 | SUBSN | subtilisin (EC 3.4 |
| 42 | 222 | 9.9 | 601 | 2 | JC4576 | serine proteinase |
| 43 | 219 | 9.7 | 384 | 2 | JC4802 | alkaline proteinase |
| 44 | 219 | 9.7 | 623 | 2 | S75376 | hypothetical prote |
| 45 | 218 | 9.7 | 361 | 2 | JH0778 | subtilisin (EC 3.4 |

ALIGNMENTS

RESULT 1

T18279
multidrug resistance transport protein - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18279
R:Shaulsky, G.; Loomis, W.F.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z18855
A:Accession: T18279
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1743 <SHA>
A:Cross-references: EMBL:U60086; NID:g1399914; PID:g1399915; PIDN:AAB03331.1
C:Genetics:
A:Gene: tagc

| | | | | | | |
|-----------------------|-------|--|-------------------------------------|------------------|------------------|---------------------|
| Query Match | 22.2% | Score | 499 | DB 2 | Length | 1743 |
| Best Local Similarity | 27.0% | Pred. No. | 3.9e-25 | | | |
| Matches | 162 | Conservative | 79 | Mismatches | 165 | Indels 194; Gaps 23 |
| QY | 11 | DVAQNN---- | FGLVGGGIVAVADTGLDTCR----- | NDS----- | SMHEAFRGKITALYAL | 56 |
| DB | 302 | DTLVNNDVRIPLRGKQIUSIADTGLDGHCFSDSKYPIPLNSVNLNR-KVTVIT | 360 | | | |
| QY | 57 | GRTNANDPNHGHPHAGSVLGN----- | ATNKGMAQANLVFQSDSGGGLGGL-- | 106 | | |
| DB | 361 | STSDSDKVDGCHGTHICGSAAGTPESSVNISSPSGLATDAKIAF---- | FDLASGSSSLTP | 417 | | |
| QY | 107 | PANLQTLFSQAYSAGARIHNSWCA---- | PVNGAYTTDSRVDVYRKN-DMTILFAAGN | 161 | | |
| DB | 418 | PSDLKQLYQPLVDAGARVHCDWSGVSVEGYTGSYSDTASIDDFLTFHDFILRAAGN | 477 | | | |
| QY | 162 | EGPGSGTISADGTAKNAITVGATENLR----- | PSFGSYADNI----- | 198 | | |
| DB | 478 | NEQVLSLLT-QSTAKVITVGAHOTIHNVLTDGPNINYQSSVDINQELICDFDSRYCN | 536 | | | |
| QY | 199 | ----- | NHVAQSSRGPTDGRKIP | 217 | | |
| DB | 537 | YTTAQCCLLESNATTGLASCCPTLLRKSVIDAANTQPLLXNENNICFSFGKPTHDGRMKP | 596 | | | |
| QY | 218 | DVMAPTGVILSARSSLA----- | PDSSFWANHDSKYAMGOTSMATIVAGNAQLAREH- | 270 | | |
| DB | 597 | ALVAPGEVITSARANGANTTDCCGGSL-PNTNALLA-IGSTNATSFARAAATILRAQYL | 654 | | | |
| QY | 271 | ----- | FVKNRGVTPKPSLLKAALJAGA----- | ADVGLGFPNGN----- | 303 | |
| DB | 655 | VDGYPTGSIVESNKLQPTGSLKALMINNAQLLNGTFLQITSSSITYPSQNFENFAGA | 714 | | | |
| QY | 304 | ---QWGRVLDKSLNVAFAVNETS----- | ----- | PLSTS | 329 | |
| DB | 715 | SLVQGWGAIRMSNMLHVNNNNNNNNNTSGITKDFDGIIGGLDLRLVKPNQWKEESLSTG | 774 | | | |

```

Qy      330 QKATYSET-----AQAGKPLK---ISLVNSDAPGSGTTASLTILVNDLDIRITAPNGTKY 379
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      775 QNTSYCFETKPSSSSSSGNNIPRVATLVMTDPPSYAGAKFNLVNNLDLTM-----IYY 829
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy      380 VGNDFTAPYDN-----NWDGRNNVENVFINAPQSGTYTVEQAVNVVPSQTES 428
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      830 RDNGSTIFYNQGSGSLGLAPTQDTLNNVEGIVHNTEPMTYRFMVAGTNVPMGPNFS 889
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
T18267
multidrug resistance protein - slime mold (Dictyostelium discoideum)
C/Species: Dictyostelium discoideum
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T18267
R:Shaullsky, G.; Kuspa, A.; Loomis, W.F.
submitted to the EMBL Data Library, January 1995
A/Description: An MDR transporter/serine protease gene is required for prestalk speciali
A/Reference number: Z18850
A/Accession: T18267
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1905 <SHA>
A/Cross-references: EMBL:U20432; NID:G664839; PID:G664840; PTDN:AAA62212.1
C:Genetics:
A:Gene: tagB

```

RESULT 3
B83891
intracellular alkaline serine proteinase aprX [imported] - Bacillus halodurans (strain
C/Species: Bacillus halodurans
C/Date: 01-Dec-2000 #sequence 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: B83891
R.;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A/Reference number: AB3650, MUID:20512582; PMID:11058132
A/Accession: B83891
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-444 <STO>
A/Cross-references: GB:AP001513; GB:BA000004; NID:gl0174345; PIDN:BAB05649.1; GSPDB:GNO
A/Experimental source: strain C-125
C/Genetics:
A/Gene: aprX

RESULT 4
A69587
intracellular alkaline serine proteinase aprX - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: A69587
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallie,
A.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.;
Koch, J.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinio,
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinio,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee,
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli,
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Ser,
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A:Authors: Yoshikawa, H.F.; Zumstoss, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: A69587
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

A;Residues: 1-442 <KUN>
A;Cross-references: GB:Z99113; GB:AL009126; NID:g2634090; PIDN:CAB13610.1; PID:ell83385;
A;Experimental source: strain 168
C;Genetics:
A;Gene: aprX
C;Superfamily: subtilisin homology
F;146-398/Domain: subtilisin homology <SBT>

Query Match 14.1%; Score 316.5; DB 2; Length 442;
Best Local Similarity 30.9%; Pred. No. 9.5e-14;
Matches 104; Conservative 41; Mismatches 99; Indels 93; Gaps 14;
QY 11 DVAQNNFLGQGGQIVAVADTGLDGRNDSMHEAFRGKITALVALGRTNNAN-----D 64
DB 136 EVVRNGQTLTGKVTVAVDGTG-----YHPDLEGR-----IGFADWVNGKTEPYD 183
QY 65 PNGHGVHAGSVGLGNATN-----KGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSAAYS 119
DB 184 DNGHGVHAGSVGLGNATN-----KGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSAAYS 119
QY 120 AGARHTNSWGAAPVN-----GAVTDSRVNVDVYKND-----MTIL 156
DB 233 -----IEGVETCIQYNEPDPEIDIMSLGDLALRYDHEQEDPLVRAVEAWAGAGIVVC 288
QY 157 FAAGNEGPGSGTISAPGAKNAITVGNATNLRPSGVSADNINVAQFSSRGPTRDGRIR 216
DB 289 VAAGNSGPDQSITASPGVSEKVIITVGLDNDNTA-----SSDDDTVASPSSRGPTVYGEK 344
QY 217 PYWAPGVYILSARSLAPDSF-----WANHDSKYAVNGTSMATPIVAGNVAQLSEHF 271
DB 345 PDILAPGVNIISLRS-----PNSVIDKLQSSRVGSGVFTMSGTSWATPICAGIALILIQ-- 399
QY 272 VKNRGVTPK--PSLLKAALIAGAADVGLGFNGNGQW 306
DB 400 -QNPDLTPDVKELK-----NGTEKW 420

RESULT 5

A41341
Microbial serine proteinase (EC 3.4.21.-), minor (vpr), precursor - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 17-Jul-1992 #sequence revision 17-Jul-1992 #text_change 20-Jun-2000
C;Accession: A41341, B41341, E39700; D69730
R;Sloma, A.; Ruto Jr., G.A.; Theriault, K.A.; Dwyer, M.; Wilson, S.W.; Pero, J.
J. Bacteriol. 173, 6889-6895, 1991
A;Title: Cloning and characterization of the gene for an additional extracellular serine
A;Reference number: A41341; MUID:92041574; PMID:1938892
A;Accession: A41341
A;Molecule type: DNA
A;Residues: 1-806 <SLO>
A;Cross-references: GB:N76590; NID:g143819; PIDN:AAA22881.1; PID:g143820
A;Accession: B41341
A;Molecule type: protein
A;Residues: 161-195 <SL2>
R;Glaeser, P.; Kunst, F.; Arnaud, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu, A.; Rapoport, G.; Danchin, A.
Mol. Microbiol. 10, 371-384, 1993
A;Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region f
A;Reference number: S39655; MUID:95020537; PMID:7934828
A;Accession: S39700
A;Molecule type: DNA
A;Residues: 1-806 <GLA>
A;Cross-references: EMBL:X73124; NID:g413923; PIDN:CAA51601.1; PID:g580871
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maude, S.; Maueel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetalle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,

A;Authors: Schleich, S.; Schröter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.; Bacillus subtilis.
A;Title: The complete genome sequence of the Gram-positive bacterium
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: D69730
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-806 <KUN>
A;Cross-references: GB:Z99123; GB:AL009126; NID:g2636240; PIDN:CAB15835.1; PID:g2636344
A;Experimental source: strain 168
C;Comment: The amino terminal sequence of the mature protein and a molecular weight of
C;Genetics:
A;Gene: vpr
A;Start codon: TTG
C;Superfamily: microbial serine proteinase vpr; subtilisin homology
C;Keywords: hydrolase; serine proteinase
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-160/Domain: propeptide #status predicted <PRO>
F;180-548/Domain: subtilisin homology #status atypical <SBT>

Query Match 13.7%; Score 309; DB 2; Length 806;
Best Local Similarity 23.2%; Pred. No. 6.8e-13;
Matches 132; Conservative 58; Mismatches 168; Indels 212; Gaps 19;
QY 16 NFGYGGQIVAVADTGLDTR-----NDSSMHEAFRGKITALVALGR 58
DB 175 DLGVTGKIKVAIDTGVENHPDLKNFGQYGVDFVNDYDPKETPTG----- 224
QY 59 TYNANDPNG-----HGTHVAGSVGLGNATNKGMAPOANLVFQSIMDSGGGLGGLPANLQTLF 114
DB 225 -----DPRGEATDHGTHVAGTVAANGTIKGVADPADATLLAYRVLPGG--SGTENVIAVG 277
QY 115 SQAYSAGARIHTNSWGAAPVNGAYTTDSRVNVDVYKNDMTILFAAGNEGPGSGTISAPCT 174
DB 278 ERAVDGADVMMNLGLNSLNNDPWATSTAL--DWANSEGVAVTNSGNSGPNGWTVGSPCT 336
QY 175 AKNAVITGATE-----NLRFSPGSY----- 194
DB 337 SREASVGTATQLPNEYAVYTFGSSYSSAKVNGYKEDDKALANKEVELVEAGIGEAKOPE 396
QY 195 -----ADNI----- 198
DB 397 GKDLTGKVAIVKGGSTAFVDKADNKAAGVAVVYNNLSGRIEANVPGMSVPTIKLSLE 456
QY 199 -----NHVAFSSRGPTRD--GRIKPDVWAPGTIYLSAR 230
DB 457 DGEKLVSAKAGETKTTFKLTVSKALGEQVADFSRGPVMDTMMIKPDISAPGVNIVSTI 516
QY 231 SSLAPDSSFWANHDSKYAVNGTSMATPIVAGNVAQLREHFVKNRGVTPKPSL--LKAAL 288
DB 517 PTHDPD-----HPYGYGSKQGTSMASPHIAGAVAVIKQ-----AKPKWSVEQIKAAI 563
QY 289 IAGAADV-----GLGFPNGNQGRVLTDLKSLNVAFVNETSPLSTOKATYVFTAAQKPL 344
DB 564 MNTAVTLKDSGDGEVYPHNAQAG-----SARINNAIKADSLVSPGSYSY---GTPL 611
QY 345 KISLVNSDAPGSTASITLVNDLDLVITAPNGTKYVGNDFAPYDNNWDG--RNNVENVF 402
DB 612 KEN-----GNETKNET-----FTENQSS--IRKSYTLEYSPNGSGISGTSRNV 655
QY 403 INAPQSGTITVEQAVNVVSPQTFSLAIV 432
DB 656 IPAHQTKATAKAVKNTKTKAQTVEGTVI 685

RESULT 6

T28159
Pyrolysin (EC 3.4.-.-) - Pyrococcus furiosus
C;Species: Pyrococcus furiosus
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 24-Oct-2000
C;Accession: 128159

R;Voorthorst, W.G.B.; Eegen, R.I.L.; Geerling, A.C.M.; Platteuw, C.; Siezen, R.J.; Vos, J. Biol. Chem. 271, 20426-20431, 1996
 A;Title: Isolation and characterization of the hyperthermostable serine protease, pyrolysin
 A;Reference number: Z20481; MUID:96355370; PMID:8702780
 A;Accession: T28159
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-1398 <VOO>
 A;Cross-references: EMBL:U55835; NID:gl556462; PID:gl556463; PIDN:AAB09761.1
 A;Experimental source: DSM3638
 C;Genetics:
 A;Gene: pls
 C;Keywords: hydrolase; serine proteinase

Query Match 13.5%; Score 304; DB 2; Length 1398;
 Best Local Similarity 26.1%; Pred. No. 3.1e-12;
 Matches 138; Conservative 54; Mismatches 176; Indels 160; Gaps 20;
 QY 21 GQGIIVAVADTGLDGTGRNDS-----SMHEAFRGKITALYALGRTNNANDPN-- 66
 DB 301 GNGYDIAYVDLTDYDFTEVPLGQYNTYDVAVFSYYVGLN--YVLAEL-----DPNGE 354
 QY 67 -----GHGTHVAGSVLGNATN-----KG 84
 DB 355 YAVEFGDGHGTHVAGTAVGYSNNDAWDLSMYSGEWEVFSRLYGMWDYTNVTDTVQG 414
 QY 85 MAPQANLVFOSIMDSGGGLGLPANLQTLFSAQAYSAGARIHTNSWG--APVNGAYTTDSR 142
 DB 415 VAPGAQIMAIRLVRSQG--RGSNWDIIIEGWTYAAATHGADVISMGLGNAPYLDGTDPEVS 472
 QY 143 NVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAITVGTATENLRPSGVAD----- 196
 DB 473 AVDELTEKYGWVFVIAAGNEGPGINIVSGPGVATKAITVGAAG--VPINVGTVVSQALGYP 531
 QY 197 -----NINHVAFSSSGPRDRGRIKPDVNAPTGYLSARSLAPDSSFWANHD 244
 DB 532 DYGYFYFPAYTV--RIAPFSSRGPRIDGEIKENVVAPGYYISLPMWIGADF----- 585
 QY 245 SKYAYMGTSMATPIVAGNVAQLREHFVKRGTVPKPSLLKAALTAGAA-----DVGLG 298
 DB 586 -----MSGTSMATPHVSGVALLISG--PKPEGIYVNPDIKKVLESATWLEGDPYTGQK 639
 QY 299 FPGNCGQGVTLDKSLNVAFVNETSPLTSQKATSFYTAQAGKPLKISLVNSDAPGSTT 358
 DB 640 YTELQDQGHVNVTKSWEI-----LKAINGTTLPIDVHWADKSYSDP 681
 QY 359 ASLTLVNDLDELVTAPNG-----TKYVGN-----FTAPYDNNWDG-----RXN 397
 DB 682 AEVLGV-DVIRGLVARNISPDIVEMWHIKYVGDTEYRTFEIYATPEWIKPFVSGSVLENN 740
 QY 398 VENVP-----INAPQSGTY-----TVEVQAYNVVPSPQTF 428
 DB 741 TEFVLRVKYDVEGLEPGLVGRRIIDPTTPVTEDEILNTIVIPERKPT 789

RESULT 7
 G83753
 C;Species: type proteinase (EC 3.4.21.-) vpr precursor [similarity] - Bacillus halodurans
 A;Title: Isolation and characterization of the hyperthermostable serine protease, pyrolysin
 C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C;Accession: G83753
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A;Reference number: A83650; MUID:20512582; PMID:11058132
 A;Accession: G83753
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-799 <STO>
 A;Cross-references: GB:AP001510; GB:BA000004; NID:gl0173440; PIDN:BA04550.1; GSPDB:GN00
 A;Experimental source: strain C-125
 C;Genetics:
 A;Gene: vpr

C;Superfamily: microbial serine proteinase vpr; subtilisin homology
 C;Keywords: hydrolase; serine proteinase
 F;1-29/Domain: signal sequence #status predicted <SIG>

Query Match 12.6%; Score 283.5; DB 2; Length 799;
 Best Local Similarity 24.5%; Pred. No. 3.4e-11;
 Matches 118; Conservative 48; Mismatches 127; Indels 189; Gaps 16;

QY 18 GLYGQGIIVAVADTGLDGTGRNDSMEHEAFRGKITALYALGR-----TNN-----AN 63
 DB 171 GYTGEGITVAILDGTVDYTHPD-----LVHAFGDKGWDFFINNDPDPQETPPG 218
 QY 64 DPNQ-----HGTHVAGSVLGNATNKGMAPOANLVFOSIMDSGGGLGLPANLQTLFSQAYS 119
 DB 219 DPGIETHGTHVAGTVAANGLIKGVAPDANLLAYRVLGPGG--RGSTAGVIAGIERAVQ 276
 QY 120 AGARIHTNSGAPNGAYTTDSRVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179
 DB 277 DGADIVNLSLGNLTNDFDFAISIAL-DWMAEGVAVTSGNSGPNNTVSGPSGSDAI 335
 QY 180 TVGATENLRPSFGSY-----ADNINH----- 200
 DB 336 SVGAT---RUPYNKYKASVETSDGIDYPSADIMGFPDSDEELLEDELGETEYAFAGLGKPG 392
 QY 201 ----- 200
 DB 393 DFEGVDVEGKIALIVRGEIPFVEKAENAKAAGAVGAIYNNVAGVQTVPEGLAIPITMLS 452
 QY 201 -----VAFSSRGPT-RDGRIPKPDVMAVPGTYILS 228
 DB 453 NEDGLKRNLENGQNTVTFSIIEFDKLVGETVADFSSRGPFVMTWIKPDVSAPGVAIVS 512
 QY 229 ARSLAPDPSFWNHSKYAYMGTSMATPIVAGNVAQLREHFVKRGTVPKPSLLKAAL 288
 DB 513 TIPHQPDDEY-----GYGSRQGTSMASPHVAGAAALLLEAH--PNMGV-----DHVKAAL 561
 QY 289 IAGADV-----GLGFPNGNGWGRVTLDKSLNVAFVNETSPLTSQKATY-SFTAQAGKP 343
 DB 562 MNTAENLVDENGNRYPHNTQAG-----SIRIVDAIESETLVPSGSHSGFTTKERKQ 615
 QY 344 LK 345
 DB 616 VE 617

RESULT 8
 S11890
 C;Species: Xanthomonas campestris pv. campestris
 A;Title: A multipurpose broad host range cloning vector and its use to characterize an
 A;Reference number: S11890; MUID:90251253; PMID:2187155
 C;Accession: S11890
 R;Jin, Y.N.; Tang, J.L.; Clarke, B.R.; Dow, J.M.; Daniels, M.J.
 Mol. Gen. Genet. 220, 433-440, 1990
 A;Title: A multipurpose broad host range cloning vector and its use to characterize an
 A;Reference number: S11890; MUID:90251253; PMID:2187155
 A;Accession: S11890
 A;Molecule type: DNA
 A;Residues: 1-580 <LU>
 A;Cross-references: EMBL:X51635; NID:948533; PIDN:CAA35962.1; PID:948534
 A;Experimental source: Xanthomonas campestris pv. campestris
 A;Note: the sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 205-Al
 C;Superfamily: subtilisin; subtilisin homology
 C;Keywords: extracellular protein; hydrolase; serine proteinase
 F;1-32/Domain: signal sequence #status predicted <SIG>
 F;168-423/Domain: subtilisin homology <SBT>

Query Match 12.4%; Score 279.5; DB 2; Length 580;
 Best Local Similarity 27.8%; Pred. No. 4.1e-11;
 Matches 133; Conservative 49; Mismatches 152; Indels 145; Gaps 26;
 QY 21 GQGIIVAVADTGL-----DTGRNDSMEHEAFRGKITALYALGRTNNAND----- 64

Db 168 GSGTVAVIDTGTSHADLNILAGYDFISDATTARDNGRGRSNAADGWDVAANECCA 227
Qy 65 -----PNGHGHVAGSVLGNATN-----KGMAPQANLV-----PQSIMD----- 98
Db 228 GIPAASSSWHGHVAGTVAATVNTTGTGAGTAYGAKVVPVVLGKCGGSLSDIADALVWA 287
Qy 99 SGGGLGLPAMLQ--TLFQSAKYSAGARIHTNSWCAPVNGAYTTDSRNVDVYVRKNDWTIL 156
Db 288 SGGTVSGIPANANPAEIVNMSLGGGSCSTTMQNA--INGAVSRGT-----TVV 334
Qy 157 FAAGNEGPG--SGTTSAPGTAKNAITVGTATEN--LRPFGSGVADNINHAQVSSRGPRDGG 213
Db 335 VAAGNDASNVSG--SLPANCANVTAATAATTSAGAKASYNFGTGI----- 377
Qy 214 RIKPDVWAPGYTILSARS--LAPDSFWANHDSKYAYMGGTSMATPIVAGNVAQLREHF 271
Db 378 ----DVSAPGSSILSTLNSGTTFGS-----ASYASNGTSMASPHVAGVVALVQS-- 424
Qy 272 VKNRGVTEK--PSLLK--AALIAGAADVGLGF-----PNGNQGWGRVTLD 312
Db 425 VAPTALTTPAAVETTLKNTARALPGACSGGCGAGIVNADAAVTAALNGSGGGGGGNTLT 484
Qy 313 KSLNVAFVNETSP--STQKATYSTTAQAKPLKISLVMSDAPGTTASLTSLVNDLDLVI- 371
Db 485 NGPTVTGLG---AATGAELNYTITVPAGSG---TLTVTTSGGG-----GDADLYVR 529
Qy 372 --TAPNGTKYVGNDFATPDYDNDGRNVENVFNAPQSGTYTVEVQAYNVPVSPQTF 428
Db 530 AGSAPTDSDAYT-----CRPYRS-----GNAETCTITAP--SGIYYVRLKAYS-----TFS 572
RESULT 9
T29090
surface layer-associated STABLE proteinase - Staphylothermus marinus
N:Alternate names: hyperthermostable proteinase
C:Species: Staphylothermus marinus
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C:Accession: T29090
R:Mayr, J.; Lupas, A.; Kellermann, J.; Eckerskorn, C.; Baumeister, W.; Peters, J.
C:Title: A hyperthermostable protease of the subtilisin family bound to the surface layer
A:Reference number: Z20559; PMID:96385442; PMID:8793300
A:Accession: T29090
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1345 <MAY>
A:Cross-references: EMBL:U57968; NID:gl374755; PID:gl374756; PIDN:AA802323.1
A:Experimental source: strain F1
C:Function:
A:Description: probably serves an exodigestive function related to the organism's energy
A>Note: stoichiometric S-layer component
Query Match 12.3%; Score 277.5; DB 2; Length 1345;
Best Local Similarity 27.3%; Pred. No. 1.7e-10;
Matches 124; Conservative 66; Mismatches 157; Indels 107; Gaps 23;
Qy 46 FRGKITALVALGRTNNDPNCHGHVHA-----GSVLGNATN-----KGMAPQANLV 92
Db 445 YQGRYLAL-----VSDFHGHTSVATVIASRGVLYDLYGDKLYRIMGVAPGAKI- 495
Qy 93 FQSIMDSGGGLGGLPANLQTLFQSAVSAG-----ARHTNSW 129
Db 496 -----AGGDWLLGNLILV--EAWLAGFNIVTEEDGYVYLSLDPFGPHRADIISNW 546
Qy 130 GA-----PVNGAYTTDSRNVDVYVRKNDWTILFAAGNEGPGSGTISAPGTAK 176
Db 547 GSTIYNFWLQPPGIDYRSSFDEILAIRNYLIGHVITVFAAGNEGPGYSNGAPGTGL 606
Qy 177 NAITVGATE--NLRSPFG--SVADNINHAQVSSRGPTDGRKIPDVMAPGYTILSARS 231
Db 607 LVITAGATLMDYTRILYGPGEFAD---EVIFFSRGPTGQGYPRPDIVN--GAFWASTR 663
Qy 232 SLAPDSFWANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAG 291

Db 664 TI--DGRGGYGAQPD-----VFGGTSEATPYTGTALVFOAYKEVYNTTDPDTAKILKSS 718
Qy 292 AADVGLGFPNGNQGWGRVTLDKSLNVAFVNE-----TSPLSTSQKATY--SFTAQAKPL 344
Db 719 AKDI--WYPAFSQSGSRVDALKAAADVFISEWLAVYVSEGIQBAFLENTYTFDGPYIGYVL 776
Qy 345 KLSLVNSD-----ARGST--TASLTIVNDLVLVIRAPNGTKYVGNDFATPYNNNDGRNV 398
Db 777 PY--LADTYYGVVYKFGSGKNFTLIVNGVAGVLSAWNTVLY--KEYTV--YDGVYDYG-- 830
Qy 399 ENNVFINAPO--SGTYTVE--VQAYNVFVSPQTF 427
Db 831 -LLFLKVPKYAYSGADYVEVVVQLENMTYPPGVF 863
RESULT 10
S23407
subtilisin (EC 3.4.21.62) 1 precursor - Bacillus sp. (strain TA39)
C:Species: Bacillus sp.
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 18-Jun-1999
C:Accession: S23407
R:Narinx, E.; Davail, S.; Feller, G.; Gerday, C.
C:Title: Biochem. Biophys. Acta 1131, 111-113, 1992
A:Title: Nucleotide and derived amino acid sequence of the subtilisin from the antarctic
A:Reference number: S23407; PMID:92256481; PMID:1581352
A:Accession: S23407
A:Molecule type: DNA
A:Residues: 1-420 <NAP>
A:Cross-references: EMBL:X62369; NID:g40200; PIDN:CAA44227.1; PID:g40201
C:Genetics:
C:Gene: sub1
C:Superfamily: subtilisin; subtilisin homology
C:Keywords: extracellular protein; hydrolase; serine proteinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-111/Domain: propeptide #status predicted <PRO>
F:112-420/Domain: microbial serine proteinase #status predicted <MAT>
F:136-374/Domain: subtilisin homology <SB1>
F:145,185,360/Active site: Asp, His, Ser #status predicted
Query Match 11.8%; Score 266; DB 1; Length 420;
Best Local Similarity 31.4%; Pred. No. 2.1e-10;
Matches 97; Conservative 40; Mismatches 114; Indels 58; Gaps 17;
Qy 21 GQQIVAVADTGLDTCNRNDSSMHEAPRGKITAL--YALGRT--NNANDPNCHGHVHAGS 75
Db 136 GGGINIAVLDTGVNTN-----HPDLRNNVEQCKDFTVGTYYTNNSCITDROGHGTHVAGS 189
Qy 76 VL-----GNATKNGMAPOANL-VFQSIMDSGGGLG-GLPANLQTLFQSAVSAGARIHTN- 127
Db 190 ALADGGTNGV-YGVAPDADLWAYKVLGGDGGYADDIAAARHAGDQATALTNTKVIWN 248
Qy 128 SWGAPVNGAYTTDSRNVDVYVRKNDWTILFAAGNEGPGSGTISAPGTAKNAITVGTATENL 187
Db 249 SLGSSGESSLIITNAVN---YSYKNGVLIITAAAGNSGPGYQSGICYPGALVNAVAVAALEN- 304
Qy 188 RPSFGSVADNINHAQVSSRGPT-RDG-----RIKPDVWAPGYTILSARSLAPDSFWA 241
Db 305 KVENGYT-----RVADFSRGYSWTGDYVAIOKGDVEISAPGAAYST-----W- 348
Qy 242 NHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSL-----LKAALIAGAA 293
Db 349 -FDGVTATISGTSMASPHAAGLAAKIWAQYPSASNVDRGELQYRAYENDILSGYVAGY 407
Qy 294 D---VGLGF 299
Db 408 DDFASGPGF 416
RESULT 11
S25935
subtilisin (EC 3.4.21.62) precursor - Bacillus sp. (strain TA41)
C:Species: Bacillus sp.

C>Date: 22-Nov-1993 #sequence_revision 20-Feb-1995 #text_change 18-Jun-1999
C/Accession: S25835
R/Davall, S.; Feiler, G.; Narinx, E.; Gerday, C.
Gene 119, 143-144, 1992
A/Title: Sequence of the subtilisin-encoding gene from an antarctic psychrotroph Bacillus
A/Reference number: S25835; MUID:93012966; PMID:1398082
A/Accession: S25835
A/Molecule type: DNA
A/Residues: 1-419 <DAV>
A/Cross-references: EMBL:X63333; NID:940199; PIDN:CAA45096.1; PID:G40199
C/Superfamily: subtilisin; subtilisin homology
C/Keywords: extracellular protein; hydrolase; serine proteinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-110/Domain: propeptide #status predicted <PRO>
F:111-419/Product: microbial serine proteinase #status predicted <MAT>
F:135-373/Domain: subtilisin homology <SBT>
F:144,184,359/Active site: Asp, His, Ser #status predicted

Query Match 11.8%; Score 265; DB 1; Length 419;
Best Local Similarity 30.9%; Pred. No. 2.4e-10;
Matches 95; Conservative 39; Mismatches 119; Indels 54; Gaps 15;

QY 21 GGGIIVAVADTGLDGRNDSSMH-EAFRGKITALYALGRT---NNANDPNHGHTHVAGSV 76
DB 135 GAGINIAVLDTGVNHPDLSXNVQCKD-----FTVGTNFTDNCCTDRQGHGTHVAGSA 189
QY 77 LGNATN----KGMAPQANL-VFQSIMDSGGGLG-GLPANLQTLFQAYSAGARIHTN-SW 129
DB 190 LANGGTGSGVGYVAPEADLWAKVLGDGSGYADDAIEAIRHAGDQATALNTKVVINWSL 249
QY 130 GAPVNGAYTDSRNVDDYKRNMDILFAAGNEGSGGTISAPGTAKNAITVGTENLRP 189
DB 250 GSGGESSLIT--NAVDAIDYKGVIIIAAGNSGPKPGSIGYPGALVNAVAVALENTIQ 306
QY 190 SFGSYADNINHVAQFSRRGPTRDG-----RIKPDVMAPGTGYILSARSLAPDSFWANH 243
DB 307 N-GTY-----RVADFSSRGHKTAGDYVIQKQDVEISAPGAAYST-----W--F 348
QY 244 DSKYAMGTSMATPIVAGNVAQLREHFVNGRVTPKPSLLKAALI-----AGAAD- 294
DB 349 DGGYATISGTSMASPHAAGLAAKIWAQSPAASNVDPVRSLOTRASVNDILSGNSAGSGDD 408
QY 295 --VGLGF 299
DB 409 IASGFGP 415

RESULT 12
JC4908
alkaline serine proteinase (EC 3.4.-.-) I precursor - Alteromonas sp.
N/Alternate names: subtilase
C/Species: Alteromonas sp.
C/Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 08-Oct-1999
C/Accession: JC4908
R/Sujiho, H.; Miyamoto, K.; Tanaka, K.; Kaidzu, Y.; Imada, C.; Okami, Y.; Inamori, Y.
Biosci. Biotechnol. Biochem. 60, 1284-1288, 1996
A/Title: Cloning and sequence analysis of a protease-encoding gene from the marine bacterium
A/Reference number: JC4908; MUID:97141200; PMID:8987544
A/Accession: JC4908
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-715 <TSU>
A/Cross-references: DDBJ:D38600; NID:G1536787; PIDN:BAAL8912.1; PID:G21602
A/Experimental source: strain O-7
C/Comment: This enzyme belongs to class I subtilisin-like family. It is a chelator-sensit
C/Genetics:
A/Genes: aprI
C/Superfamily: subtilisin homology
C/Keywords: hydrolase
F:1-40/Domain: signal sequence #status predicted <SIG>
F:41-150/Domain: amino-terminal propeptide #status predicted <ATP>
F:151-496/Product: alkaline serine protease I #status predicted <MAT>
F:182-452/Domain: subtilisin homology <SBT>

F:497-715/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F:239-294,335-372,478-481/Disulfide bonds: #status predicted

Query Match 11.7%; Score 262.5; DB 2; Length 715;
Best Local Similarity 25.3%; Pred. No. 7.4e-10;
Matches 118; Conservative 51; Mismatches 176; Indels 121; Gaps 20;

QY 21 GGGIIVAVADTGLDGRNDSSMH-EAFRGKITALYALGRTNNAND- 65
DB 182 GGGVAVADTGYRPHLDLDAILPGYDMISNTVFANDGGAEDNDARDPGDAVTRGCGT 241
QY 66 -----NGHGTHTVAGSVLGNATN----KGMAPQANLVPFQSIMDSGGGLGLP 107
DB 242 DSSGQVPVPRADQDSWHGTHVAGTAAVTNNGEGVAGVADAKVVPFVRL--GKGGGLT 298
QY 108 ANLQTLFQAYSAGARIHTNSGAPV-----NGAYTDSRNVDDYKRNMDTILFAAG 160
DB 299 SDIADGIWASGGSDRVPANANPANNVMSLGGGACSAATTQAINQARNNTVIVIAAG 358
QY 161 NEFGSGGTISAPGTAKNAITVGAT--ENLRPSFGSYADNINHVAQFSRRGPTRDGRIPD 218
DB 359 NDNDNSANYN-PGNCNGVVNVASVGRDGRSRYVSNYGANI-----D 398
QY 219 VMAPGTGYILSARSLAPDSFWANHDS-----KYAMGTSMATPIVAGNVAOLREH 270
DB 399 VAAPG---GAQSFADDPGILSTHNSGSGAPSNDSYHSQGTSMAPHVAGVAAALIK- 453
QY 271 FVKNRGVTPKPS-----LLK--AALTAGAADVGLGFPNGQNGRVTLDKSLNAVAV 320
DB 454 -----AKPSATPDEVETILKNTTFSFAGSC-----SNCGTGVVDAANAANALG 497
QY 321 N-ETSPILSTSQATYFTAQAGKPLKISLVWSDAP-GSTTASLTLL---VNDLDLVI---T 372
DB 498 DVVTPPTGTLEDGVAKTGLSGAAGSNQFTFDVPAGKTNVFTTMSGGTGADLIYKLG 557
QY 373 APNGTKVGVNDFAPYDNNWGDGRNNVENVFINAPQSGTYTVEVQAY 418
DB 558 QPTSSSVY-----DCRPYEG-----GNAEVCSEFAPQAGTYHVINGY 594

RESULT 13

D75393

serine proteinase, subtilase family - Deinococcus radiodurans (strain R1)

C/Species: Deinococcus radiodurans

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C/Accession: D75393

R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A/Reference number: A75250; MUID:20036896; PMID:10567266

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-627 <WHI>

A/Cross-references: GB:AE001990; GB:AE000513; NID:G6459214; PIDN:AAF11026.1; PID:G64592

A/Experimental source: strain R1

C/Genetics:

A/Genes: DR1459

A/Map position: 1

Query Match 11.6%; Score 261.5; DB 2; Length 627;

Best Local Similarity 28.0%; Pred. No. 7.3e-10;

Matches 92; Conservative 49; Mismatches 122; Indels 65; Gaps 14;

QY 18 GLYGGQIIVAVADTGLDGRNDSSMH-EAFRGKITALYALGR-----TNNANDPNHGHTHV 72

DB 161 GFKQGNTRIGHLSDGIDPS-----HPELAGKVAARQFENGEGDRVSSQPHDTTDTGHT 214

QY 73 AGSVLGNATKGMAPQANLVPFQSIM-DSGGGLGLPANLQTLFQAYS----GARIHTN 127

DB 215 AGLLVGSKV--GVAPGAKVTSALVLPNNEGTFQVAGMQYVLDPPONNADTDGDADVNM 272

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 7.49423 Seconds
(without alignments)
3008.498 Million cell updates/sec

Title: US-09-985-689A-7

Perfect score: 2252

Sequence: 1 NDVARGIVKADVAQNNGFLY.....EVQAYNVFSPQTFSLAIVH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 499 | 22.2 | 1743 | 1 | TAGC D1CDI |
| 2 | 494.5 | 22.0 | 1905 | 1 | TAGB D1CDI |
| 3 | 309 | 13.7 | 806 | 1 | SUBV_BACSU |
| 4 | 306 | 13.6 | 1398 | 1 | PLS_PFRFU |
| 5 | 279.5 | 12.4 | 580 | 1 | EXPR_XANCP |
| 6 | 266 | 11.8 | 420 | 1 | SUBT_BACSU |
| 7 | 254 | 11.3 | 534 | 1 | PROA_VIBAL |
| 8 | 249 | 11.1 | 401 | 1 | THES_BACSP |
| 9 | 245.5 | 10.9 | 269 | 1 | SUBS_BACLE |
| 10 | 245.5 | 10.9 | 321 | 1 | ISP_BACCS |
| 11 | 245.5 | 10.9 | 380 | 1 | ELVA_BACAO |
| 12 | 245.5 | 10.9 | 380 | 1 | ELVA_BACCS |
| 13 | 245.5 | 10.9 | 513 | 1 | AQU1_THEAQ |
| 14 | 244.5 | 10.9 | 269 | 1 | PRTM_BACSP |
| 15 | 243.5 | 10.8 | 894 | 1 | WPRA_BACSU |
| 16 | 239.5 | 10.6 | 269 | 1 | SUBB_BACLE |
| 17 | 232 | 10.3 | 1052 | 1 | MSIP_HUMAN |
| 18 | 230 | 10.2 | 326 | 1 | ISP_PAEPO |
| 19 | 230 | 10.2 | 1167 | 1 | SCAI_STRPY |
| 20 | 229 | 10.2 | 402 | 1 | ALP_CBPAC |
| 21 | 229 | 10.2 | 1181 | 1 | SCA2_STRPY |
| 22 | 228 | 10.1 | 1052 | 1 | MSIP_CRIGR |
| 23 | 228 | 10.1 | 1052 | 1 | MSIP_MOUSE |
| 24 | 228 | 10.1 | 1052 | 1 | MSIP_RAT |
| 25 | 227.5 | 10.1 | 378 | 1 | ELVA_BACSP |
| 26 | 226 | 10.0 | 379 | 1 | SUBT_BACLI |
| 27 | 226 | 10.0 | 1433 | 1 | SUBF_BACSU |
| 28 | 225.5 | 10.0 | 319 | 1 | ISPI_BACSU |
| 29 | 225.5 | 10.0 | 404 | 1 | SMPI_HAGPO |
| 30 | 223 | 9.9 | 530 | 1 | HLV_HALI7 |
| 31 | 222 | 9.9 | 382 | 1 | SUBT_BACAM |
| 32 | 218 | 9.7 | 381 | 1 | SUBN_BACNA |
| 33 | 217 | 9.6 | 381 | 1 | SUBT_BACSA |

| | | | | | |
|----|-------|-----|------|---|------------|
| 34 | 217 | 9.6 | 381 | 1 | SUBT_BACST |
| 35 | 216.5 | 9.6 | 422 | 1 | TKSU_PFRKO |
| 36 | 216 | 9.6 | 381 | 1 | SUBT_BACSU |
| 37 | 214 | 9.5 | 275 | 1 | SUBD_BACLI |
| 38 | 214 | 9.5 | 275 | 1 | SUBT_BACPU |
| 39 | 213.5 | 9.5 | 279 | 1 | THET_THEVU |
| 40 | 210 | 9.3 | 467 | 1 | ISP6_SCHPO |
| 41 | 209.5 | 9.3 | 293 | 1 | PRTT_TRIAL |
| 42 | 208 | 9.2 | 595 | 1 | BPRX_BACNO |
| 43 | 206 | 9.1 | 1374 | 1 | YQ86_CABEL |
| 44 | 203.5 | 9.0 | 409 | 1 | ALP_TRIHA |
| 45 | 203.5 | 9.0 | 603 | 1 | BPRV_BACNO |

ALIGNMENTS

RESULT 1

| ID | TAGC D1CDI | STANDARD; | PRT; 1743 AA. |
|----|--|-----------|---------------|
| AC | Q23868; | | |
| DT | 01-NOV-1997 (Rel. 35, Created) | | |
| DT | 01-NOV-1997 (Rel. 35, Last sequence update) | | |
| DT | 15-MAR-2004 (Rel. 43, Last annotation update) | | |
| DE | Prestalk-specific protein tagC precursor (EC 3.4.21.-). | | |
| GN | TAGC. | | |
| OS | Dictyostelium discoideum (Slime mold). | | |
| OC | Eukaryota; Mycetozoa; Dictyostelid; Dictyostelium. | | |
| OX | NCBI_TaxID=44689; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=AX4; | | |
| RX | MEDLINE=97140317; PubMed=8986798; | | |
| RA | Shaulsky G., Escalante R., Loomis W.F.; | | |
| RT | "Developmental signal transduction pathways uncovered by genetic suppressors." | | |
| RL | Proc Natl Acad Sci U.S.A. 93:15260-15265(1996). | | |
| CC | -!- FUNCTION: Intercellular communication via tagC may mediate integration of cellular differentiation with morphogenesis (by similarity). | | |
| CC | -!- SIMILARITY: In the N-terminal section; belongs to peptidase family S8. | | |
| CC | -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY. | | |
| CC | -!- SIMILARITY: STRONG, TO TAGB. | | |
| CC | ----- | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | | |
| CC | ----- | | |
| DR | EMBL; U60086; AAB03331.1; -. | | |
| DR | PIR; T18279; T18279. | | |
| DR | DictyBase; DDB0001795; tagC. | | |
| DR | InterPro; IPR003593; AAA_Atpase | | |
| DR | InterPro; IPR011140; ABC_TM_transpt.. | | |
| DR | InterPro; IPR003439; ABC_Transporter. | | |
| DR | InterPro; IPR000209; Peptidase_S8. | | |
| DR | Pfam; PF00664; ABC membrane; 1. | | |
| DR | Pfam; PF00005; ABC tran; 1. | | |
| DR | Pfam; PF00082; Peptidase S8; 1. | | |
| DR | PRINTS; PR00723; SUBTILISIN. | | |
| DR | SMART; SM00382; AAA; 1. | | |
| DR | PROSITE; PSS0929; ABC_TMIF; 1. | | |
| DR | PROSITE; PS00211; ABC_TRANSPORTER_1; 1. | | |
| DR | PROSITE; PS00893; ABC_TRANSPORTER_2; 1. | | |
| DR | PROSITE; PS00136; SUBTILASE ASP; FALSE_NEG. | | |
| DR | PROSITE; PS00137; SUBTILASE HIS; 1. | | |
| DR | PROSITE; PS00137; SUBTILASE_SER; FALSE_NEG. | | |
| KW | Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane; | | |

| | |
|--------|-------------|
| P29142 | bacillus st |
| P58502 | pyrococcus |
| P04189 | bacillus su |
| P00781 | bacillus li |
| P07518 | bacillus pu |
| P04072 | thermoactin |
| P40903 | schizosacch |
| P20015 | tritirachiu |
| P42780 | bacteroides |
| Q09541 | caenorhabdi |
| Q03420 | trichoderma |
| P42779 | bacteroides |

| FT | SIGNAL | 1 | 31 | POTENTIAL | |
|----|-----------------------|---|--|--------------------------------------|-----|
| FT | CHAIN | 32 | 1905 | PRESTALK-SPECIFIC PROTEIN TAGS. | |
| FT | DOMAIN | 378 | 700 | PROTEASE. | |
| FT | DOMAIN | 1518 | 1756 | ABC TRANSPORTER. | |
| FT | TRANSMEM | 1011 | 1031 | POTENTIAL. | |
| FT | TRANSMEM | 1076 | 1096 | POTENTIAL. | |
| FT | TRANSMEM | 1121 | 1141 | POTENTIAL. | |
| FT | TRANSMEM | 1210 | 1230 | POTENTIAL. | |
| FT | TRANSMEM | 1309 | 1329 | POTENTIAL. | |
| FT | TRANSMEM | 1332 | 1352 | POTENTIAL. | |
| FT | ACT SITE | 387 | 387 | CHARGE RELAY SYSTEM (BY SIMILARITY). | |
| FT | ACT SITE | 432 | 432 | CHARGE RELAY SYSTEM (BY SIMILARITY). | |
| FT | ACT SITE | 695 | 695 | CHARGE RELAY SYSTEM (BY SIMILARITY). | |
| FT | NP BIND | 1553 | 1560 | ATP (POTENTIAL). | |
| FT | DOMAIN | 63 | 67 | POLY-GLN. | |
| FT | DOMAIN | 95 | 104 | POLY-ASN. | |
| FT | DOMAIN | 107 | 134 | POLY-ASN. | |
| FT | DOMAIN | 311 | 321 | POLY-SER. | |
| FT | DOMAIN | 833 | 837 | POLY-SER. | |
| FT | DOMAIN | 838 | 844 | POLY-GLY. | |
| FT | DOMAIN | 871 | 876 | POLY-LEU. | |
| FT | DOMAIN | 1012 | 1015 | POLY-ILE. | |
| FT | DOMAIN | 1386 | 1389 | POLY-GLU. | |
| FT | DOMAIN | 1398 | 1404 | POLY-GLY. | |
| FT | DOMAIN | 1445 | 1450 | POLY-ASN. | |
| FT | DOMAIN | 1765 | 1779 | POLY-ASN. | |
| FT | DOMAIN | 1782 | 1785 | POLY-SER. | |
| FT | DOMAIN | 1807 | 1812 | POLY-PRO. | |
| FT | DOMAIN | 1813 | 1860 | POLY-GLN. | |
| FT | DOMAIN | 1872 | 1878 | POLY-PRO. | |
| FT | CARBOHYD | 594 | 594 | N-LINKED (GLCNAC. . .) (POTENTIAL). | |
| FT | CARBOHYD | 621 | 621 | N-LINKED (GLCNAC. . .) (POTENTIAL). | |
| FT | CARBOHYD | 672 | 672 | N-LINKED (GLCNAC. . .) (POTENTIAL). | |
| FT | CARBOHYD | 747 | 747 | N-LINKED (GLCNAC. . .) (POTENTIAL). | |
| FT | CARBOHYD | 823 | 823 | N-LINKED (GLCNAC. . .) (POTENTIAL). | |
| FT | CARBOHYD | 1172 | 1172 | N-LINKED (GLCNAC. . .) (POTENTIAL). | |
| FT | CARBOHYD | 1522 | 1522 | N-LINKED (GLCNAC. . .) (POTENTIAL). | |
| FT | CARBOHYD | 1658 | 1658 | N-LINKED (GLCNAC. . .) (POTENTIAL). | |
| FT | SEQUENCE | 1905 AA; | 212518 MW; | B8E223FA8B9AE13C CRC64; | |
| Qy | Query Match | 22.0%; | Score 494.5; | DB 1; Length 1905; | |
| Db | Best Local Similarity | 28.1%; | Pred. No. 2e-24; | | |
| Qy | Matches 162; | Conservative | 72; | Mismatches 162; Indels 181; Gaps 22; | |
| Qy | 19 | LYGGQIVAVADTGLDTR | ---NDS-----SMHEAFRGKITALYALGRINNDP--N | 66 | |
| Db | 376 | LRRGQILSTADTGLDGSCHCFSDSKYP | IPFNQVNNHRKVTYI-----TYHDNEDYVN | 430 | |
| Qy | 67 | GHGTHVAGSVLGN | -----ATNKMAQANLVFQSIMDSGGGLGLPANLQTLFQAY | 118 | |
| Db | 431 | GHGTHVCSAGTDESDSWAISRSGLATDAKIAFYD | -LSSGSSEPTPEDYQMYKPLY | 489 | |
| Qy | 119 | SAGARIHTNSWGA | ---PVNGAYTDSRVDYVRK-NDMTILPAAGNEGSGSTISAPG | 173 | |
| Db | 490 | DAGARVHGDSWGSVSLQYGYGSDDAGGIDAFLEYEPFESILRAAGN-NELFASLLAQ | 546 | | |
| Qy | 174 | TAKNAITVCATENLRPSTGS | -----YADNI----- | 198 | |
| Db | 549 | TAKNAITVGAQTAAHVNTVSDALEYDFSDNANFORPCLFKKYCNVYATKCCSESVNK | 608 | | |
| Qy | 199 | -----NHVAQFSSRGTRDGRKIPDVMAPTGYILSARS | 231 | | |
| Db | 609 | GLQCCPASIKQNASDSTFTQPFQYNNNMGSFSSKGPTHDRKLPDIPVAPGEYITSARS | 668 | | |
| Qy | 232 | S-----LAPDSSFWANHDSKYVNGCTSMATPIVAGNVAQIREF | -----VKNR | 275 | |
| Db | 669 | NGENSTDCQDGSGL--PNANGLMSISGTSMATPLATAATTLRQLVLDGYPTGSGSEVEN | 726 | | |
| Qy | 276 | GVTPKPSLLKALIAAGADV | -----LCFPNGNCQWGRVTLDKSLN | 316 | |
| Db | 727 | KLLPTGSLIKALMINNAQLNGTYFWSASSTNFSNAIFQINGANLIQWGAIRNNWLY | 786 | | |
| Qy | 317 | VAFVNETSP | -----YSTAQ----- | AGKP | 343 |

Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.; "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*." Nature 390:249-256(1997).

CC -!- FUNCTION: NOT REQUIRED FOR GROWTH OR SPORULATION.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- PTM: PROBABLY UNDERGOES C-TERMINAL PROCESSING OR PROTEOLYSIS.

CC -!- SIMILARITY: Belongs to peptidase family S8.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M76590; AAA22881.1; -

DR EMBL; X73124; CAA51601.1; -

DR EMBL; Z93123; CAB15835.1; -

DR F1R; A41341; A41341.

DR HSP; P00782; 2SST.

DR MEROPS; S08.LPA; -

DR Subtilisin; BG10591; vpr.

DR InterPro; IPR003137; PA.

DR InterPro; IPR00209; Peptidase_S8.

DR InterPro; IPR009020; Protease_inhib.

DR Pfam; PF02225; PA; 1.

DR Pfam; PF00082; Peptidase_S8; 1.

DR PRINTS; PR00723; SUBTILISIN.

DR PROSITE; PS00136; SUBTILASE_ASP; 1.

DR PROSITE; PS00137; SUBTILASE_HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.

KW Hydrolase; Serine protease; Zymogen; Signal; Complete proteome.

FT SIGNAL 1 28

FT PROPEP 29 160

FT CHAIN 161 806

FT ACT_SITE 189 189

FT ACT_SITE 233 233

FT ACT_SITE 534 534

FT ACT_SITE 806 AA; 85608 MW; F984E3BF0B869DDD CRC64;

SQ SEQUENCE 806 AA; 85608 MW; F984E3BF0B869DDD CRC64;

Query Match 13.7%; Score 309; DB 1; Length 806;

Best local similarity 23.2%; Pred. No. 8.3e-13;

Matches 132; Conservative 58; Mismatches 168; Indels 212; Gaps 19;

QY 16 NFGLYGQGIQVAVADGLDTR-----NDSSMEAFRGKITALYALGR 58

DB 175 DLGVTGKIKVAIDTGVENHDLKGNFGQYKGYDFVNDYDKETPTG----- 224

QY 59 TNNANDPNG-----HGTHVAGSVLGNATNKGMAFOANLVFQSIKSGGLGGLPANLQTLF 114

DB 225 -----DPRGEATDHGTGTVAGTVAANGTIKGVAPDATTLLAYRVVLPGG--SGTTENVIAGV 277

QY 115 SQAYSACARTHTSWGAPVNGAYTTDSRNVDVVRKNDMTILFAAGNEGPGSGTISAPGT 174

DB 278 ERAVQDGVNVLNLSLNPDPWATSTAL-DWAMSEGVAVTNSGNSGNGPWTGSPGT 336

QY 175 AKNAITVGATE-----NLRPFGSGY----- 194

DB 337 SREAIYSGATQPLNEVAVTFGYSYSAKVMGYNKEDDVKALNKEVELVEAGIGEAQDFE 396

QY 195 -----ADNI----- 198

DB 397 GKDJTGKVVVKGSIAPVDKADNAKAGAGIMVYNNLSGETEIANVPGMSVPTIKLSLE 456

QY 199 -----NHVAQFSRRGFTED-GRIFPDVMAPGTIILSAR 230

DB 457 DGEKLVSALKAGETKTTFKLTIVSKALGEQVADFSSRGFVMDTWMIKPDISAPGVNVTI 516

QY 231 SSLAPDSFWNHDISKVAYMGSTMATPIVAGNVAQLREHFVKNRGVTPKPSL--LKAAL 288

Db 517 PTHDPD-----HPYGYGSKQGTSMASPHIAGAVAIKQ-----AKPKWSVEQIKAAI 563

QY 289 IGAADV-----GLGFPNGQGWGRVTLDKSLNVAFVNETSPSTQKATYSTAQAGKPL 344

Db 564 MNTAVTLKDSGEVYPHNAQAGAG-----SARIMNAIKADSLVSPGYSY---GTFLL 611

QY 345 KISLWSDAPGSTTASLTLLVNDLDTAPNGTKYVGNDFTAPYDNNWDG--RNNVENVF 402

Db 612 KEN-----GNETKNET-----FTIENQSS-IRKSYTILEYSPNGSGISTGTSRVV 655

QY 403 INAPQSGTVEVQVYVNPVSPQTFLSLAV 432

Db 656 IPAQQTGKATARKVNTKTKTACGYEGTVI 685

RESULT 4

PLS PYRFU

ID - PLS PYRFU STANDARD; PRT; 1398 AA.

AC P72186;

DT 16-OCT-2001 (Rel. 40, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Pyrolysin precursor (BC 3.4.21.-).

GN PLS OR PF0287.

OS Pyrococcus furiosus.

OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;

OC Pyrococcus.

OX NCBI_TaxID=2261;

EN [1]

RP SEQUENCE FROM N.A. SEQUENCE OF 150-184, AND CHARACTERIZATION.

RC STRAIN=Vci / DSM 3638 / ATCC 43587 / JCM 8422;

RA MEDLINE=96355370; PubMed=8702780;

RX Voorhorst W.G.B., Eggen R.I.L., Geerling A.C.M., Platteuw C., Siezen R.J., de Vos W.M.;

RT "Isolation and characterization of the hyperthermostable serine protease, pyrolysin, and its gene from the hyperthermophilic archaeon Pyrococcus furiosus."

RT J. Biol. Chem. 271:20426-20431(1996).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Vci / DSM 3638 / ATCC 43587 / JCM 8422;

RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;

RT "The complete sequence of the Pyrococcus furiosus genome."

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

RP CHARACTERIZATION, AND 3D-STRUCTURE MODELING.

RC MEDLINE=21079021; PubMed=11210516;

RA de Vos W.M., Voorhorst W.G.B., Dijkgraaf M., Kluskens L.D., Van der Oost J., Siezen R.J.;

RT "Purification, characterization, and molecular modeling of pyrolysin and other extracellular thermostable serine proteases from hyperthermophilic microorganisms."

RL Meth. Enzymol. 330:383-393(2001).

CC -!- FUNCTION: Has endopeptidase activity toward caseins, casein fragments including alpha-SI-casein and synthetic peptides.

CC -!- SUBCELLULAR LOCATION: Cell envelope associated.

CC -!- PTM: LWM pyrolysin seems to be produced by autoprolytic activation of HMW pyrolysin.

CC -!- PTM: Glycosylated.

CC -!- MISCELLANEOUS: Thermostable; high activity at 95 degrees Celsius.

CC -!- SIMILARITY: Belongs to peptidase family S8.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U55835; AAB09761.1; -

DR EMBL; AE010153; AAL80411.1; -

DR F1R; T28159; T28159.


```

DR EMBL; AE012184; AAM40166.1; -.
DR PIR; S11890; S11890.
DR HSP; P00782; 2SBI.
DR MRO08; S08.UFA; -.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR InterPro; IPR009020; Protease_inhib.
DR Pfam; PFO0082; Peptidase_S8; 1.
DR Pfam; PFO4151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR HydroLase; Serine protease; Zymogen; Signal; Complete proteome.
FT SIGNAL 1 32
FT PROPEP 1 32
FT CHAIN 2137 580
FT ACT_SITE 177 177
FT ACT_SITE 237 237
FT ACT_SITE 409 409
FT DISULFID 225 273
FT DISULFID 315 352
FT DISULFID 450 454
SQ SEQUENCE 580 AA; 57228 MW; 8C9A2CAE4B7F47CB CRC64;

Query Match
Best Local Similarity 12.4%; Score 279.5; DB 1; Length 580;
Matches 133; Conservative 49; Mismatches 152; Indels 145; Gaps 26;

QY 21 GQGIIVAVADTGL-----D*GRNDSMHEAFRGKITALYALGRTNNAND----- 64
DB 168 GSGTVVAVIDTGTSHADLNANILAGYDFTSDATTARDNGRDSNADGDDWYAAECGA 227

QY 65 -----PNGHGHVAGSVLGNATN-----KGMAPQANLV-----FQSIMD----- 98
DB 228 GIPAASSWHGTHVAGTAAVTNNTTGVAGTAYGAKVVPVRLGKGGSLSDIADAIWMA 287

QY 99 SGGGLGGLPANLQ--TLFSQAYSAGARIHTNSGAPVNGAYTTDSRNVDYVRKNDMTIL 156
DB 288 SGTVSGIPANANPAEVINNSLGGGSCSTTQNA--INGAVRGIT-----TVV 334

QY 157 FAAGNPGPG--SGTISAPGTAKNAITVGATEN--LRPSFGSYADNINHVAQFSRGPTRDG 213
DB 335 VAAGNDASNVSG--SLPANCANVIAVAATTSAGAKASYNPGTGI----- 377

QY 214 RIKPDVWAPCTYILSARSS--LAPDSSFWANHDSKYAYMGFTSMATPIVAGNVAQLREHF 271
DB 378 ----DVSAPGSSLSLTNSGTTTPGS-----ASTASYNGTSMASPHVAGVVALVQS-- 424

QY 272 VKNRGVTPK--PSLLK--AALIAGAADVGLGF-----PNGNQGWGRVTLD 312
DB 425 VAPTALTPAAVETLLKNTARALPGACSGCGAGIVNADAATAAINGGSGGGGGGNTLT 484

QY 313 KSLNVAFVNETSPLTSQKATYSFTAQAGKPLKISLVWSDAPGSTTASITLVNDLDLVI- 371
DB 485 NGTPVTGLG-----AATGAELNYITIVPAGSG---TLTIVTSSGGS-----GDADLYVR 529

QY 372 --TAPNGTKYVGNDFPTAPDNNWGDGRNNVENFINAPQSGCTYVEVQAYNVFVPSQTFS 428
DB 530 AGSAPTDASVT-----CRPYRS-----GNAETCTITAP-SGTYYVRLKAYS-----TFS 572

RESULT 6
SUBT_BACS9
ID SUBT_BACS9 STANDARD; PRT; 420 AA.
AC P28842;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Subtilisin precursor (EC 3.4.21.62).
GN SUBI.
OS Bacillus sp. (strain TA39).
OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.

```

```

OX NCBI_TaxID=29336;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92256481; PubMed=1581352;
RT Narinx E., Davail S., Feller G., Gerday C.;
RT "Nucleotide and derived amino acid sequence of the subtilisin from
RT the antarctic psychrotroph Bacillus TA39.";
RL Biochim. Biophys. Acta 1131:111-113(1992).
CC -!- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
CC it catalyzes the hydrolysis of proteins and peptide amides.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large uncharged residue
CC in P1. Hydrolyzes peptide amides.
CC -!- COFACTOR: Binds 1 calcium ion per subunit (Potential).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MISCELLANEOUS: Still active at temperatures close to 0 degrees
CC Celsius; it has a marked heat lability.
CC -!- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
CC sporulation, and many mutations which block sporulation at early
CC stages affect expression levels of subtilisin. However, subtilisin
CC is not necessary for normal sporulation.
CC -!- SIMILARITY: Belongs to peptidase family S8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X62369; CAA44227.1; -.
CC PIR; S23407; S23407.
CC HSP; Q99405; IMPT.
CC MEROPS; S08.UFA; -.
CC InterPro; IPR000209; Peptidase_S8.
CC Pfam; PF00082; Peptidase_S8; 1.
CC PRINTS; PR00723; SUBTILISIN.
CC PROSITE; PS00136; SUBTILASE_ASP; 1.
CC PROSITE; PS00137; SUBTILASE_HIS; 1.
CC PROSITE; PS00138; SUBTILASE_SER; 1.
CC HydroLase; Sporulation; Serine protease; Zymogen; Metal-binding;
CC Calcium-binding; Signal.
KW SIGNAL 1 26
KW PROPEP 27 111
KW CHAIN 112 420
KW ACT_SITE 145 145
KW ACT_SITE 182 182
KW ACT_SITE 360 360
KW METAL 115 115
KW METAL 154 154
SQ SEQUENCE 420 AA; 44086 MW; AE4F121BD32B26EC CRC64;

Query Match
Best Local Similarity 11.8%; Score 266; DB 1; Length 420;
Best Local Similarity 31.4%; Pred. No. 2.3e-10;
Matches 97; Conservative 40; Mismatches 114; Indels 58; Gaps 17;

```

```

QY 21 GQGIIVAVADTGLDTRNDSMHEAFRGKITAL--VALGRT---NNANDPNGHGHVAGS 75
DB 136 GGGINIAVLDTGVNTN-----HPDLNNVVEQKDFVTGVTYNNSTDRQGHGTHVAGS 189

QY 76 VL-----GNATNKGMAPQANL-VFQSIMDSGGGLG-GLPANLOTLPQAYSAGARIHTN- 127
DB 190 ALADGGTGNGV-YGVAPDADLWAYKVLGDDGSGYADDAIAAIRHAGDQATALNTKVYINM 248

QY 128 SWGAPVNGAVTDSRNVDVVRKNDMTILFAAGNEGPGSGTISAPGTAKNAITVGATENL 187
DB 249 SLGSGGESSLITNAVN---YSYNGKVLIIIAAGNSGPGYQSGIPALVNAVVALEN- 304

QY 188 RPSFGSYADNINHVAQFSRGPRT-RDG-----RIKPDVWAPGTYYILSARSSLAPDSFWA 241
DB 305 KVENGTY-----RVADPSSRGYSWTDGVAIQKGDVEISAPGAATYST-----W- 348

```


QY 242 NHDSKYAYMGTSMTPIVAGNVAQLREHFVKRGVTPKPSL-----LKAALIAGAA 293
 Db 349 -FDGGYATISGTSMSFHAAGLAAKIWAQIPASNVVDVREGELQYRAYENDILSGYYAGY 407
 QY 294 D---VGLGF 299
 Db 408 DDFASGFGF 416

RESULT 7
 ID PROA_VIBAL STANDARD; PRT; 534 AA.
 AC P16588;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Alkaline serine exoprotease A precursor (EC 3.4.21.-).
 GN PROA.
 OS *Vibrio alginolyticus*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=663;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89326126; PubMed=2546861;
 RA Deane S.M., Robb F.T., Robb S.M., Woods D.R.;
 RT "Nucleotide sequence of the *Vibrio alginolyticus* calcium-dependent,
 RT detergent-resistant alkaline serine exoprotease A.";
 RL Gene 76:281-288(1989).
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M25499; AAA27550.1; -
 DR F1R; J50173; J50173.
 DR HSP; Q99405; 1MPT.
 DR MEROPS; S08.050; -
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR007250; PPC.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydrolase; Serine protease; Zymogen; Signal.
 FT SIGNAL 1 21
 FT PROPEP 22 141
 FT CHAIN 142 534
 FT ACT_SITE 180 180
 FT ACT_SITE 213 213
 FT ACT_SITE 363 363
 SQ SEQUENCE 534 AA; 55930 MW; 84B96D9C649D4226 CRC64;
 Query Match 11.3%; Score 254; DB 1; Length 534;
 Best Local Similarity 26.1%; Pred. No. 1.9e-09;
 Matches 124; Conservative 63; Mismatches 138; Indels 150; Gaps 26;
 QY 7 IVKADVAQNN--FGL-----YQGOIVAVADTGLDTRGDSMHEAF 46
 Db 137 IVSADANQINAIWGLDRIDQRNLPIDNNYSANFDGTGTAYIDTGV-----NNAHYEF 190
 QY 47 RKGITALY-ALGRITNANDPNHGTHVAGSVLGNATNKGMAPQANLVFQSIW--DSGGGL 103
 Db 191 GGRSVSGYDFVDNDADASDCNGHGHVAGTIGSSL--YGVAKVNLVGVRLVSCSGSGST 248

QY 104 GGLPANLQTLPSQAYSAGARIHNSWGAIPVNGVYTTDSRWVDYVR---KNDMTILFAAG 160
 Db 249 SGVIAGDWTVAANA--SGPSVANMSLGGQGSVA-----LDSAVSAGVSGVSFMLAAG 299
 QY 161 NEGFSGSTISAPGTAKNAITVGAT--ENLRPSFGSVADNINHYAQPSSRGPTDRGKRPD 218
 Db 300 NSNADACNYS-PARVATGVTVGSTTTDARSSFSNWGSCV-----D 339
 QY 219 VMAPGTIILSARSLAPDSSFANHDSKYAYMGTSMTPIVAGNVAQLREHFVKRGV 278
 Db 340 VFAPGSQIKSA-----W-YDGGYKTIISGTSMTAPHVAG-VAAL--YLQENS SVS 384
 QY 279 PKPSLLKAAIAGAADVGLGFPNGNQMGVTLDK-SLVAVFVNETS----- 324
 Db 385 --PSQVEALLIVSRAT-----GKVTTRGSVKNLLYSLDADCGQCGGPDPT 430
 QY 325 -----PL-----STSQKATYSFATAQAKPLKISLVNSDAPGSGTASLTLYNDLD 368
 Db 431 PDPEGKLTSGVPVSGLSGSGQVAYVYVDVEAGRLTVQMY-----GGS-----GDAD 478
 QY 369 LVITAPNGTKYVGNDFTFAPYDNNNDGR-----NNVENVFVINAPOSCTTVEVOAYN 419
 Db 479 LYLRF--GAK-----PTLNAWDCRPFYKGNNTCTVTSATOSGRHYHVMIOGYS 523

RESULT 8
 ID THES_BACSP STANDARD; PRT; 401 AA.
 AC Q45670;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Thermophilic serine proteinase precursor (EC 3.4.21.-) (Ak.1
 DE protease).
 OS *Bacillus* sp. (strain AK1).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1409;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95085262; PubMed=7993087;
 RA Maciver B., McHale R.H., Saul D.J., Bergquist P.L.;
 RT "Cloning and sequencing of a serine proteinase gene from a
 RT thermophilic *Bacillus* species and its expression in *Escherichia*
 RT coli.";
 RL Appl. Environ. Microbiol. 60:3981-3988(1994).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.80 ANGSTROMS) OF 122-401.
 RX MEDLINE=20057863; PubMed=10588904;
 RA Smith C.A., Toogood H.S., Baker H.M., Daniel R.M., Baker E.N.;
 RT "Calcium-mediated thermostability in the subtilisin superfamily: the
 RT crystal structure of *Bacillus* Ak.1 protease at 1.8-A resolution.";
 RL J. Mol. Biol. 294:1027-1040(1999).
 CC -1- COFACTOR: Binds 3 calcium ions and 1 sodium ion per subunit.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: Has a pH optimum of 8.5, a temperature optimum of
 CC 75 degrees Celsius.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L29506; AAA3688.1; -
 DR PIR; I39974; I39974.
 DR PDB; 1DBI; 18-NOV-99.
 DR MEROPS; S08.009; -
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR005020; Protease_inhib.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR

DR PRINTS; PRO0723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE SER; 1.
 KW Hydrolases; Serine protease; Zymogen; Metal-binding; Calcium-binding;
 FT SIGNAL; 3D-structure.
 FT SIGNAL 1 24
 FT PROPEP 25 121
 FT CHAIN 122 401
 FT ACT_SITE 160 160
 FT ACT_SITE 193 193
 FT ACT_SITE 347 347
 FT METAL 126 126
 FT METAL 168 168
 FT METAL 169 169
 FT METAL 171 171
 FT METAL 179 179
 FT METAL 184 184
 FT METAL 186 186
 FT METAL 204 204
 FT METAL 207 207
 FT METAL 209 209
 FT METAL 211 211
 FT METAL 297 297
 FT METAL 300 300
 FT METAL 323 323
 FT DISULFID 258 260
 FT TURN 127 128
 FT HELIX 129 132
 FT TURN 135 136
 FT HELIX 137 139
 FT TURN 140 140
 FT HELIX 141 144
 FT TURN 145 147
 FT TURN 152 153
 FT STRAND 155 160
 FT TURN 165 166
 FT TURN 168 173
 FT STRAND 174 179
 FT TURN 180 183
 FT STRAND 184 184
 FT HELIX 193 202
 FT STRAND 219 224
 FT TURN 228 229
 FT HELIX 234 246
 FT TURN 247 248
 FT STRAND 251 254
 FT HELIX 263 274
 FT TURN 275 276
 FT STRAND 278 282
 FT STRAND 285 285
 FT TURN 298 299
 FT STRAND 301 306
 FT TURN 308 309
 FT STRAND 312 312
 FT TURN 314 315
 FT STRAND 316 316
 FT TURN 320 321
 FT STRAND 324 327
 FT STRAND 331 335
 FT TURN 336 338
 FT STRAND 339 343
 FT HELIX 346 362
 FT TURN 363 364
 FT HELIX 367 376
 FT TURN 377 377
 FT STRAND 379 379
 FT TURN 382 383
 FT STRAND 384 384
 FT TURN 385 387
 FT STRAND 388 388
 FT TURN 391 392

FT HELIX 395 399
 FT TURN 400 401
 SQ SEQUENCE 401 AA; 42835 WS; 1C736EP4A89F256P CRC64;
 Query Match 11.1%; Score 249; DB 1; Length 401;
 Best Local Similarity 31.7%; Pred. No. 2.8e-09;
 Matches 85; Conservative 29; Mismatches 96; Indels 58; Gaps 12;
 QY 14 QNNFGLY-----GQQQIVAVADTGLDTRNDSSNHEAFRGKITALYALGNTNNDP- 65
 DB 137 QNTYTDYANDVTKSGSGQIAVIDGVD-----YTHPLDGKVIKGYDF--VDNDYDPM 188
 QY 66 --NGHGTHTVAG---SVLGNATN-KGVAPOANLVFQSIMDSGGGLGLPANLQTLFSQAYS 119
 DB 189 DLANNHGTHTVAGTAAETNNATGIAGMAPNTRILAVALDRNG--SCTUSDIADAIYAAD 246
 QY 120 AGARIHTNSWGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179
 DB 247 SGAEVINLSGC---DCHTTTLENAYVYAWKGSVVVAAGNN--GSSTTEPASVENVI 301
 QY 180 TVGATENLRPFSGYADNINHVAQPSRQPTRDGRIKPDVMAPGYVILSARSLAPDSF 239
 DB 302 AVGA-----VDQYDRLASFNSYGTW-----VDVAPGVDIVSTITG----- 337
 QY 240 WANHDSKYAYMGTSMTATPIVAGNVAQL 267
 DB 338 -----NRYAYMSTGTSNAPSHVAGLAALL 360
 RESULT 9
 SUBS_BACLE STANDARD; PRT; 269 AA.
 AC P29600;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DS Subtilisin Savinase (EC 3.4.21.62) (Alkaline protease).
 OS Bacillus lentus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI_TaxID=1467;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).
 RX MEDLINE=92148829; PubMed=1738156;
 RA Betzel C., Klupsch S., Papendorf G., Hastrup S., Branner S.,
 Wilson K.S.;
 RT "Crystal structure of the alkaline proteinase Savinase from Bacillus
 lentus at 1.4-A resolution."
 RL J. Mol. Biol. 223:427-445(1992).
 RN [2]
 RP STRUCTURE BY NMR.
 RX MEDLINE=96184541; PubMed=8654411;
 RA Remerowski M.L., Pepermans H.A.M., Hilbers C.W., van de Ven F.J.M.;
 RT "Backbone dynamics of the 269-residue protease Savinase determined
 from 15N-NMR relaxation measurements."
 RL Eur. J. Biochem. 235:629-640(1996).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (0.78 ANGSTROMS).
 RX MEDLINE=98426039; PubMed=9753430;
 RA Kuhn P., Knapp M., Soltis S.M., Ganshaw G., Thoene M., Bott R.;
 RT "The 0.78-A structure of a serine protease: Bacillus lentus
 subtilisin."
 RL Biochemistry 37:13446-13452(1998).
 CC -!- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
 it catalyzes the hydrolysis of proteins and peptide amides.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
 for peptide bonds, and a preference for a large uncharged residue
 in P1. Hydrolyzes peptide amides.
 CC -!- COFACTOR: Binds 2 calcium ions per subunit.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- BIOTECHNOLOGY: Used as a detergent protease. Sold under the name
 Savinase by Novozymes.
 CC -!- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
 sporulation, and many mutations which block sporulation at early

stages affect expression levels of subtilisin. However, subtilisin is not necessary for normal sporulation.

-|- SIMILARITY: Belongs to peptidase family S8.

PDB; 1C9J; 06-OCT-99.

PDB; 1C9N; 10-JAN-01.

PDB; 1GCI; 11-NOV-98.

PDB; 1IAV; 11-JUL-01.

PDB; 1JEA; 26-NOV-97.

PDB; 1SVN; 14-OCT-96.

MEROPS; S08.103; ..

InterPro; IPR000209; Peptidase_S8.

Pfam; PF00082; Peptidase_S8; 1.

PRINTS; PR00723; SUBTILISIN.

PROSITE; PS00136; SUBTILASE ASP; 1.

PROSITE; PS00137; SUBTILASE HIS; 1.

PROSITE; PS00138; SUBTILASE_SER; 1.

Hydrolase; Sporulation; Serine protease; Metal-binding;

Calcium-binding; 3D-structure.

ACT SITE 32 32 CHARGE RELAY SYSTEM.

ACT SITE 62 62 CHARGE RELAY SYSTEM.

ACT SITE 215 215 CHARGE RELAY SYSTEM.

METAL 2 2 CALCIUM 1.

METAL 40 40 CALCIUM 1.

METAL 73 73 CALCIUM 1.

METAL 75 75 CALCIUM 1.

METAL 77 77 CALCIUM 1.

METAL 79 79 CALCIUM 1.

METAL 163 163 CALCIUM 1.

METAL 165 165 CALCIUM 2.

METAL 166 166 CALCIUM 2.

STRAND 2 2 CALCIUM 2.

HELIX 6 10 (VIA CARBONYL OXYGEN).

TURN 11 12 (VIA CARBONYL OXYGEN).

TURN 13 18 (VIA CARBONYL OXYGEN).

TURN 19 20 (VIA CARBONYL OXYGEN).

TURN 24 25 (VIA CARBONYL OXYGEN).

TURN 27 32 (VIA CARBONYL OXYGEN).

TURN 39 40 (VIA CARBONYL OXYGEN).

TURN 43 48 (VIA CARBONYL OXYGEN).

TURN 51 52 (VIA CARBONYL OXYGEN).

TURN 62 71 (VIA CARBONYL OXYGEN).

TURN 78 78 (VIA CARBONYL OXYGEN).

TURN 84 85 (VIA CARBONYL OXYGEN).

TURN 87 92 (VIA CARBONYL OXYGEN).

TURN 94 94 (VIA CARBONYL OXYGEN).

TURN 96 97 (VIA CARBONYL OXYGEN).

STRAND 100 100 (VIA CARBONYL OXYGEN).

STRAND 102 114 (VIA CARBONYL OXYGEN).

TURN 115 116 (VIA CARBONYL OXYGEN).

STRAND 119 122 (VIA CARBONYL OXYGEN).

STRAND 126 126 (VIA CARBONYL OXYGEN).

TURN 131 142 (VIA CARBONYL OXYGEN).

TURN 133 144 (VIA CARBONYL OXYGEN).

STRAND 146 150 (VIA CARBONYL OXYGEN).

STRAND 161 161 (VIA CARBONYL OXYGEN).

TURN 162 164 (VIA CARBONYL OXYGEN).

TURN 166 167 (VIA CARBONYL OXYGEN).

STRAND 168 174 (VIA CARBONYL OXYGEN).

TURN 176 177 (VIA CARBONYL OXYGEN).

STRAND 180 180 (VIA CARBONYL OXYGEN).

TURN 182 183 (VIA CARBONYL OXYGEN).

TURN 188 189 (VIA CARBONYL OXYGEN).

STRAND 190 195 (VIA CARBONYL OXYGEN).

STRAND 199 203 (VIA CARBONYL OXYGEN).

TURN 204 206 (VIA CARBONYL OXYGEN).

TURN 207 211 (VIA CARBONYL OXYGEN).

STRAND 214 231 (VIA CARBONYL OXYGEN).

TURN 233 234 (VIA CARBONYL OXYGEN).

HELIX 237 246 (VIA CARBONYL OXYGEN).

TURN 247 247 (VIA CARBONYL OXYGEN).

STRAND 249 249 (VIA CARBONYL OXYGEN).

HELIX 254 254 (VIA CARBONYL OXYGEN).

TURN 254 254 (VIA CARBONYL OXYGEN).

TURN 254 254 (VIA CARBONYL OXYGEN).

FT TURN 258 259
FT STRAND 261 261
FT HELIX 264 267
FT TURN 268 269
SQ SEQUENCE 269 AA; 26698 MW; 4D89F8778999BF8D CRC64;

Query Match 10.9%; Score 245.5; DB 1; Length 269;

Best Local Similarity 31.0%; Pred. No. 2.8e-09;

Matches 88; Conservative 32; Mismatches 99; Indels 65; Gaps 13;

QY 8 VKADVAQNFGLYGQGIQVAVADTGLDTRNDSSMHEAFRGKITALYALGRTNNDPNG 67

DB 11 VQAPAAHNR-GLTSGVKVAVLDTGIST-----HPDLNIRGASGVFGEPE-STQDNG 61

QY 68 HGTHVAGSV--LGNATN-KMAPCANLVFQSIINDSGGLGGLPANLQTLFQAYSAGARI 124

DB 62 HGTHVAGTTAALNNISGLVGAAPSALYAVKVLGASG--SGSVSSIAQGLEWAGNNGMHV 119

QY 125 HTNSWAGVNGAYTTDSRVNDDYVRKNDMTIIFAGNEGSGSTISAPGTAKNAITVGAT 184

DB 120 ANLSLGSFSPSATLEQAVN---SATSRGLVVAASGNSGAGS--ISYPARYANAMAVGAT 174

QY 185 E-NLRPSFGSIYADNINHVAFQSSRGFTDGRKIPDVPAPGVYILARSLSLAPDSSFWAN 242

DB 175 DQNNRASFQYAGL-----DIVAPGVNVQSTPG----- 205

QY 243 HDSKYAYMGGTSMATPIVAGNVA-----OLREHFVKV 274

DB 206 --STYASLNGTSMATHVAGAAALVKQKNPSNVTQIRNH-LKN 246

RESULT 10

ISP_BACCS

ID ISP_BACCS STANDARD; PRT; 321 AA.

AC P29140;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Intracellular alkaline protease (EC 3.4.21.-).

OS Bacillus clausii.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=79890;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=221 / ATCC 21522 / JCM 9139 / DSM 2512;

RA Kato C., Nakano Y., Yamamoto M., Horikoshi K.;

RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.

CC -|- SIMILARITY: Belongs to peptidase family S8.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;

CC or send an email to license@isb-sib.ch).

CC -----

CC ENBL; D10730; BAA01573.1; -

DR PIR; S27501; S27501.

DR HSSP; P00782; IS01.

DR MEROPS; S08.030; -

DR InterPro; IPR000209; Peptidase_S8.

DR Pfam; PF00082; Peptidase_S8; 1.

DR PRINTS; PR00723; SUBTILISIN.

DR PROSITE; PS00136; SUBTILASE ASP; 1.

DR PROSITE; PS00137; SUBTILASE HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.

DR Hydrolase; Serine protease.

KW Hydrolyase; Serine protease.

FT ACT_SITE 49 49 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 86 86 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT_SITE 250 250 CHARGE RELAY SYSTEM (BY SIMILARITY).

SQ SEQUENCE 321 AA; 33747 MW; 621168D9F1044026 CRC64;

```

Query Match      10.9%; Score 245.5; DB 1; Length 321;
Best Local Similarity 30.4%; Pred. No. 3.5e-09;
Matches 94; Conservative 40; Mismatches 114; Indels 61; Gaps 15;

QY 21 GGGQIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTNANDPNHGTHVAGSVLQNA 80
DB 40 GAGQIIGVDTGQVDDHPDLAERI:GGVNLTTDYG-GVETNPSDNNGHGTHTVAGTAAAE 98
QY 81 TNK--GMAPQANLVFQSIM--DSGGGLGGLPANLQTLFSQAYSAG--ARHTNSWGAPV 133
DB 99 TCGVVGVAPKADLFIKALSDGGSGGEMGWIKAIRYAVDWRGPKGEQWRITWLSLGP- 157
QY 134 NGAYTDSNRVD---YVRKNDMTILFAAGNEGPG---SGTISAPETAKNAITVGTAB-N 186
DB 158 -----TDSSELHDAVKYASNNVSVVCAAGNEGDEPTNEPAYPAAYNEVIAVAVD 212
QY 187 LRFPSGVSADNHNHVAQFSRGTDRGRIKPDVMAFGTVILSARSLAPDSFWANHDSK 246
DB 213 LR-----LSDF-----PNTNEEI--DIVAPGVGIKSTYL-----DSG 242
QY 247 YAYMGTSMTPIVAGNVAQL-----REHFVKVRGVTTPKPSLLKAALIAAGADVGLGFPNG 302
DB 243 YAEISGTSMAHPVAGALALIINLANDAFKRTLSETE----ICAQLVRRATPI--GFTAQ 296
QY 303 NOGNGWRVTL 311
DB 297 DKONGFLTL 305

RESULT 11
ELVA_BACAO STANDARD; PRT; 380 AA.
ID ELVA_BACAO
AC P27693;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alkaline protease precursor (EC 3.4.21.-)
OS Bacillus alcalophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1445;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PB92;
RX MEDLINE=91282483; PubMed=2059048;
RA van der Laan J.C., Gerritse G., Mulleners L.J.M., van der Hoek R.A.,
RA Quax W.J.;
RT "Cloning, characterization, and multiple chromosomal integration of a
RT Bacillus alkaline protease gene."
RL Appl. Environ. Microbiol. 57:901-909(1991).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
RC STRAIN=PB92;
RX MEDLINE=92390330; PubMed=1518788;
RA van der Laan J.C., Tepljakov A.V., Kelders H., Kalk K.H., Misset O.,
RA Mulleners L.J.M., Dijkstra B.W.;
RT "Crystal structure of the high-alkaline serine protease PB92 from
RT Bacillus alcalophilus."
RL Protein Eng. 5:405-411(1992).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
RX MEDLINE=93078250; PubMed=1447775;
RA Sobek H., Hecht H.-J., Ahle W., Schomburg D.;
RT "X-ray structure determination and comparison of two crystal forms of
RT a variant (Asn15Arg) of the alkaline protease from Bacillus
RT alcalophilus refined at 1.85-A resolution."
RL J. Mol. Biol. 228:108-117(1992).
RN [4]
RP STRUCTURE BY NMR OF 112-380.
RC STRAIN=PB92;
RX MEDLINE=97277237; PubMed=9115441;
RA Martin J.R., Mulder F.A., Karimi-Nejad Y., van der Zwan J.,
RA Mariani M., Schipper D., Boelens R.;
RT "The solution structure of serine protease PB92 from Bacillus

```

```

RT alcalophilus presents a rigid fold with a flexible substrate-binding
RT site."
RL Structure 5:521-532(1997).
CC -|- COPACTOR: Binds 2 calcium ions per subunit.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- SIMILARITY: Belongs to peptidase family S8.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M65086; AAA22212.1; -
DR EMBL; A13738; CAA01128.1; -
DR PIR; A49778; A49778.
DR PDB; 1AH2; 15-APR-98.
DR MEROPS; S08.038; -.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR009020; Protease_inhib.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR Hydrolase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
KW Signal; 3D-structure.
FT SIGNAL 1 27 POTENTIAL.
FT PROPEP 28 112
FT CHAIN 113 380 ALKALINE PROTEASE.
FT ACT_SITE 143 143 CHARGE RELAY SYSTEM.
FT ACT_SITE 173 173 CHARGE RELAY SYSTEM.
FT ACT_SITE 326 326 CHARGE RELAY SYSTEM.
FT METAL 113 113 CALCIUM 1.
FT METAL 151 151 CALCIUM 1.
FT METAL 184 184 CALCIUM 1 (VIA CARBONYL OXYGEN).
FT METAL 186 186 CALCIUM 1.
FT METAL 188 188 CALCIUM 1 (VIA CARBONYL OXYGEN).
FT METAL 190 190 CALCIUM 1 (VIA CARBONYL OXYGEN).
FT METAL 274 274 CALCIUM 2 (VIA CARBONYL OXYGEN).
FT METAL 276 276 CALCIUM 2 (VIA CARBONYL OXYGEN).
FT METAL 279 279 CALCIUM 2 (VIA CARBONYL OXYGEN).
FT HELIX 121 120 TURN
FT TURN 121 123
FT HELIX 125 130
FT TURN 131 131
FT TURN 135 136
FT STRAND 138 142
FT TURN 150 151
FT STRAND 156 157
FT HELIX 173 182
FT STRAND 198 201
FT TURN 207 208
FT HELIX 213 226
FT STRAND 230 233
FT HELIX 242 252
FT TURN 253 255
FT STRAND 257 258
FT TURN 267 268
FT TURN 272 275
FT STRAND 279 279
FT STRAND 282 285
FT STRAND 291 291
FT TURN 299 300
FT STRAND 303 306
FT STRAND 310 314
FT TURN 315 317
FT STRAND 318 322
FT HELIX 325 342
FT TURN 344 345
FT HELIX 348 358

```

```
FT HELIX 375 380
SQ SEQUENCE 380 AA; 38853 MW; 539EA72771B6682C CRC64;

Query Match 10.9%; Score 245.5; DB 1; Length 380;
Best Local Similarity 31.0%; Pred. No. 4.4e-09;
Matches 88; Conservative 32; Mismatches 99; Indels 65; Gaps 13;

QY 8 VKADVAQNNGFLYGQCIIVAVADTGLDTRGNDSSMHEAFRGKITALYALGRTNANDPNG 67
DQ 122 VQAPAAHNR-GLTGSQVKVAVLDTGIST-----HPDLNIRGGSFVPGEP-STQDNG 172
QY 68 HGTHVAGSV--LGNATN-KGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSAGARI 124
DQ 173 HGTHVAGTTAALNNSTIGVLGAPNAELVAVKVLGASG--SGSVSSIAQGLEWAGNNGMHV 230
QY 125 HTNSWGPVNGAYTTDSRVDDVVRKNDMTILPAAGNEGPGSGTISAPGTAKNAITVGAT 184
DQ 231 ANLSLGSPPSPATLEQAVN---SATSRGVLVVAASGNSGAGS--ISYFARYANAMAVGAT 285
QY 185 E--NLRFPSGYSADNINHVAFSSRGPTDRGRIKPDVMAFPYILSARSSLAPDSSFAN 242
DQ 286 DOWNREASPSQYAGL-----DIVAPGVNVQSTYPG----- 316
QY 243 HDSKYAYMGTSVATPIVAGNVA-----QLREHFVN 274
DQ 317 --STYASLNGTSWATPHVAGAAALVKQKPNPSWNSVQIRNH-LKN 357

RESULT 12
ELVA_BACCS STANDARD; PRT; 380 AA.
AC P41362;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alkaline protease precursor (EC 3.4.21.-).
OS Bacillus clausii.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=79880;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=221 / ATCC 21522 / JCM 9139 / DSM 2512;
RX MEDLINE=93043753; PubMed=1368952;
RA Takami H., Kobayashi T., Kobayashi M., Yamamoto M., Nakamura S.,
RA Acno R., Horikoshi K.;
RT "Molecular cloning, nucleotide sequence, and expression of the
RT structural gene for alkaline serine protease from alkaliphilic
RT Bacillus sp. 221.";
RL Biosci. Biotechnol. Biochem. 56:1455-1460 (1992).
RN [2]
SEQUENCE OF 112-129.
RC STRAIN=221 / ATCC 21522 / JCM 9139 / DSM 2512;
RA Horikoshi K.;
RL (In) Horikoshi K. (eds.);
RL Microorganisms in alkaline environments, pp.187-194, VCH,
RL Weinheim (1991).
CC -/- COFACTOR: Binds 2 calcium ions per subunit (By similarity).
CC -/- SUBCELLULAR LOCATION: Secreted.
CC -/- SIMILARITY: Belongs to peptidase family S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; S48754; AAC60420.1; -
CC DR ENBL; D13157; BAA02442.1; -
CC DR ENBL; A26817; CAA01836.1; -
CC DR ENBL; A22550; CAA01611.1; -
CC HSP; P29600; IGC1.
```

```
DR MEROPS: S08.103; -
DR InterPro: IPR000209; Peptidase S8.
DR InterPro: IPR009020; Protease_Inhib.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PRO0723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE ASP; 1.
DR PROSITE: PS00137; SUBTILASE HIS; 1.
DR PROSITE: PS00138; SUBTILASE SER; 1.
DR Hydrolase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
KW Signal.
FT SIGNAL 1 27 POTENTIAL.
FT PROPEP 28 111
FT CHAIN 112 380
FT ACT_SITE 143 143 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 173 173 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 326 326 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT METAL 113 113 CALCIUM 1 (BY SIMILARITY).
FT METAL 151 151 CALCIUM 1 (BY SIMILARITY).
FT METAL 184 184 CALCIUM 1 (VIA CARBONYL OXYGEN)
(BY SIMILARITY).
FT METAL 186 186 CALCIUM 1 (BY SIMILARITY).
FT METAL 188 188 CALCIUM 1 (VIA CARBONYL OXYGEN)
(BY SIMILARITY).
FT METAL 190 190 CALCIUM 1 (VIA CARBONYL OXYGEN)
(BY SIMILARITY).
FT METAL 274 274 CALCIUM 2 (VIA CARBONYL OXYGEN)
(BY SIMILARITY).
FT METAL 276 276 CALCIUM 2 (VIA CARBONYL OXYGEN)
(BY SIMILARITY).
FT METAL 279 279 CALCIUM 2 (VIA CARBONYL OXYGEN)
(BY SIMILARITY).
FT SEQUENCE 380 AA; 38826 MW; 5F73ABC68D5B6831 CRC64;

Query Match 10.9%; Score 245.5; DB 1; Length 380;
Best Local Similarity 31.0%; Pred. No. 4.4e-09;
Matches 88; Conservative 32; Mismatches 99; Indels 65; Gaps 13;

QY 8 VKADVAQNNGFLYGQCIIVAVADTGLDTRGNDSSMHEAFRGKITALYALGRTNANDPNG 67
DQ 122 VQAPAAHNR-GLTGSQVKVAVLDTGIST-----HPDLNIRGGSFVPGEP-STQDNG 172
QY 68 HGTHVAGSV--LGNATN-KGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSAGARI 124
DQ 173 HGTHVAGTTAALNNSTIGVLGAPNAELVAVKVLGASG--SGSVSSIAQGLEWAGNNGMHV 230
QY 125 HTNSWGPVNGAYTTDSRVDDVVRKNDMTILPAAGNEGPGSGTISAPGTAKNAITVGAT 184
DQ 231 ANLSLGSPPSPATLEQAVN---SATSRGVLVVAASGNSGAGS--ISYFARYANAMAVGAT 285
QY 185 E--NLRFPSGYSADNINHVAFSSRGPTDRGRIKPDVMAFPYILSARSSLAPDSSFAN 242
DQ 286 DOWNREASPSQYAGL-----DIVAPGVNVQSTYPG----- 316
QY 243 HDSKYAYMGTSVATPIVAGNVA-----QLREHFVN 274
DQ 317 --STYASLNGTSWATPHVAGAAALVKQKPNPSWNSVQIRNH-LKN 357

RESULT 13
AQLI_THEAQ STANDARD; PRT; 513 AA.
AC P08594;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aqualysin I precursor (EC 3.4.21.-).
GN PSI1.
OS Thermus aquaticus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=271;
RN [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 15-23.
RP
```

RC STRAIN-YT1;
RX MEDLINE=90216674; PubMed=2182621;
RA Terada I., Kwon S.-T., Miyata Y., Matsuzawa H., Ohta T.;
RT "Unique precursor structure of an extracellular protease, aqualysin
I, with NH2- and COOH-terminal pro-sequences and its processing in
Escherichia coli.";
RT J. Biol. Chem. 265:6576-6581(1990).
RN [2]
RP SEQUENCE OF 75-442 FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-YT1;
RX MEDLINE=88225062; PubMed=3286255;
RA Kwon S.-T., Terada I., Matsuzawa H., Ohta T.;
RT "Nucleotide sequence of the gene for aqualysin I (a thermophilic
alkaline serine protease) of *Thermus aquaticus* YT-1 and
characteristics of the deduced primary structure of the enzyme.";
RT Eur. J. Biochem. 173:491-497(1988).
RN [3]
RP SEQUENCE OF 128-170.
RC MEDLINE=88151937; PubMed=3162211;
RA Matsuzawa H., Tokugawa K., Hamaoki M., Mizoguchi M., Taguchi H.,
Terada I., Kwon S.-T., Ohta T.;
RT "Purification and characterization of aqualysin I (a thermophilic
alkaline serine protease) produced by *Thermus aquaticus* YT-1.";
RT Eur. J. Biochem. 171:441-447(1988).
RL -1- FUNCTION: Aqualysin I is a thermophilic alkaline serine protease.
CC The optimal temperature for its caseinolytic activity is 80
CC degrees Celsius.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DEVELOPMENTAL STAGE: Secreted from the early stationary phase
CC until the time the cells cease to grow.
CC -1- PTM: THE N- AND C-TERMINAL PRO-SEQUENCES ARE REMOVED THROUGH THE
CC PROTEOLYTIC ACTIVITY OF AQUALYSIN I ITSELF, IN THAT ORDER. THE
CC C-TERMINAL PRO-SEQUENCE IS REQUIRED FOR TRANSLOCATION OF THE
CC PROTEASES ACROSS THE OUTER MEMBRANE.
CC -1- PTM: Two disulfide bonds are present.
CC -1- SIMILARITY: Belongs to peptidase family S8.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90108; BAA14135.1; -;
DR EMBL; X07734; CAA30559.1; -;
DR PIR; A35742; A35742.
DR HSP; P06873; 2PRK.
DR MEROPS; S08.051; -;
DR InterPro; IPR000209; Peptidase S8.
DR InterPro; IPR003020; Protease Inhib.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
KW Hydrolase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 14
FT PROPEP 15 127
FT CHAIN 128 408 AQUALYSIN I.
FT PROPEP 409 513
FT ACT_SITE 166 166 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 197 197 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 349 349 CHARGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 513 AA; 53913 MW; DDFDFE6D4A50B795 CRC64;
Query Match 10.9%; Score 245.5; DB 1; Length 513;
Best Local Similarity 26.0%; Pred. No. 6.5e-09;
Matches 121; Conservative 43; Mismatches 145; Indels 157; Gaps 24;
7 IVKADVAQNNFGLV---GGQIVAVADTGLDTGRNDSSMEAFRGKITALLY-ALGRITNNA 62
QY | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

Db 140 IDQDLPLSNSYTYTATGRGVNVVVDITGIRT-----THREFGGRARVGYDALG--GNG 191
Qy 63 NDENGHGTHVAGSVLGNATNKGMAPANLVFQSIMD--SGGGLGLPANLQTLFSAQYSA 120
Db 192 QDCNHHGTHVAGTIGG--VTYGVAKAVNLYAVRVLDCNGSGSTSGVIAGVDWV----- 242
Qy 121 GARIHTN-----SWGAPVNGAYTDSRVNVDYVRKN---DMTILFAAGNEGPGSGTISA 171
Db 243 -TENHREPAVANWSLGGGVSTA-----LDNAVKSIAAGVVVYAAGNDNANACNYS- 293
Qy 172 PGAKNAITVGAT--ENLRPSFGSYADNHNHVAQFSRGPTRDGRKPDVMAPGTVILSA 229
Db 294 PARVAEALTVCATSSDARASFNYGSCV-----DLFAPGASIPSA 334
Qy 230 RSSLAPDSEFWANEDSKYAYMGGTSMATPIVAGNVAQLREHFVKRQGVTP---KPSLLKA 286
Db 335 -----WYTSDDTATQTLNGTSMATPHVAG-VAAL--YLEQNPSATPASVASAILNG 381
Qy 287 ALIAGAADVGLGFPNGQNGRVTLDKSLNVAFWNETSPLSTSOXATSYFTAQAGKPLKI 346
Db 382 ATTGRLSGIGSGSPN-----TASLTIVNDLDELVTAPNGTKY-----RLLYS- 401
Qy 347 SLVMSDAPGST-----TASLTIVNDLDELVTAPNGTKY-----RLLYS- 383
Db 402 --LLSSGSGTAPCTSCSYVTGSLGPGDYNF---QPNGTYYYSYPAGTHRAWLRGPAGTD 456
Qy 384 FTAPYDNNWDGRN-----NVNMFNAPQSGTYTVEVOAYN 419
Db 457 FDL-YLRWDGSRWLTVGSGTGTSBESLSYSTAGYLMRIYAYS 501
RESULT 14
ID PRM_BACSP STANDARD; PRM; 265 AA.
AC Q99405;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE M-protease (EC 3.4.21.-).
OS *Bacillus* sp. (strain KSM-K16).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RA Yanane T., Kani T., Hatanaka T., Suzuki A., Ashida T.,
RA Kobayashi T., Ito S., Yamashita O.;
RT "Structure of a new alkaline serine protease (M-protease) from
Bacillus sp. KSM-K16.";
RL Acta Crystallogr. D 51:199-206(1995).
RN [2]
RP SEQUENCE OF 1-23; AND CHARACTERIZATION.
RX MEDLINE=9535832; PubMed=7632397;
RA Kobayashi T., Hakanada Y., Adachi S., Hitomi J., Yoshimatsu T.,
RA Koike K., Kawai S., Ito S.;
RT "Purification and properties of an alkaline protease from
alkalophilic *Bacillus* sp. KSM-K16.";
RL Appl. Microbiol. Biotechnol. 43:473-481(1995).
CC -1- COFACTOR: Binds 2 calcium ions per subunit.
CC -1- SUBUNIT: Monomer.
CC -1- SIMILARITY: Belongs to peptidase family S8.
DR PDB; 1MPT; 22-JUN-94.
DR MEROPS; S08.010; -;
DR InterPro; IPR000209; Peptidase S8.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
KW Hydrolase; Serine protease; Metal-binding; Calcium-binding;
3D-structure.
KW ACT_SITE 32 32 CHARGE RELAY SYSTEM.
FT ACT_SITE 62 62 CHARGE RELAY SYSTEM.
FT ACT_SITE 215 215 CHARGE RELAY SYSTEM.

```

FT METAL 2 2 FT METAL 1. CALCIUM 1.
FT METAL 40 40 FT METAL 1. CALCIUM 1.
FT METAL 73 73 FT METAL 1. (VIA CARBONYL OXYGEN).
FT METAL 75 75 FT METAL 1.
FT METAL 77 77 FT METAL 1. (VIA CARBONYL OXYGEN).
FT METAL 79 79 FT METAL 1. (VIA CARBONYL OXYGEN).
FT METAL 163 163 FT METAL 2 (VIA CARBONYL OXYGEN).
FT METAL 165 165 FT METAL 2 (VIA CARBONYL OXYGEN).
FT METAL 168 168 FT METAL 2 (VIA CARBONYL OXYGEN).
FT METAL 6 10 FT METAL 2 (VIA CARBONYL OXYGEN).
FT HELIX 11 12
FT TURN 12 12
FT HELIX 13 18
FT TURN 19 20
FT TURN 24 25
FT TURN 27 32
FT STRAND 27 32
FT TURN 39 40
FT STRAND 43 48
FT TURN 51 52
FT HELIX 52 71
FT TURN 62 71
FT STRAND 87 92
FT TURN 96 97
FT TURN 102 114
FT TURN 115 116
FT STRAND 119 122
FT STRAND 126 126
FT HELIX 131 142
FT TURN 143 144
FT STRAND 146 150
FT STRAND 161 161
FT TURN 162 164
FT TURN 166 167
FT STRAND 169 174
FT TURN 176 177
FT STRAND 180 180
FT TURN 182 183
FT TURN 188 189
FT STRAND 192 195
FT STRAND 199 203
FT TURN 204 206
FT STRAND 207 211
FT HELIX 214 231
FT TURN 233 234
FT HELIX 237 246
FT TURN 247 247
FT STRAND 249 249
FT HELIX 254 257
FT TURN 258 259
FT STRAND 261 261
FT HELIX 264 267
FT TURN 268 269
SQ SEQUENCE 269 AA; 26723 MW; 7A03C86D534A1D07 CRC64;

Query Match 10.9%; Score 244.5; DB 1; Length 269;
Best Local Similarity 31.0%; Pred. No. 3.2e-09;
Matches 88; Conservative 32; Mismatches 99; Indels 65; Gaps 13;

QY 8 VRADVAQNNGFLYGQCIVAVADTGLDTGRNDSMHEAFRGKITAYALGRTNANDPNG 67
Db 11 VQAPAHNR-GLTSGVKVAVLDTGST-----HPDLNIRGGASFVPEGP-STQDNG 61
QY 68 HGTHVAGSV--IGNATN-KGMAPQANLVFQSIWDSGGGLGGLPANLQTLFQAYVSAGARI 124
Db 62 HGTHVAGTIALNNSIGVLGVAPSALVAVKVLGASG--SGSVSSIAQGLEWAGNNGMHV 119
QY 125 HTNSWCAPVNGAYTTDSRNVDYVRKNDWTLTILFAAGNPGSGGTISAPGTAKNATTVGAT 184
Db 120 ANLSLGSPPSPTLEQAVN---SATSRGVLVVVAAGNSGAGS--ISYFARYANAVAVGAT 174
QY 185 E-NLRPSFGSVADNINHVAFQSSRGPTDRGIRKPDVWAPGYTILSARSLAPDSSFWAN 242
Db 175 DQNNRNASPSQYAGI-----DIVAPGVNVQSTYFG----- 205

```

```

QY 243 HDSKYAYMGTSMTATPIVAGNVA-----QLREHFVK 274
Db 206 --STYASLNGTSMATPHVAGVAALVKQKNPSWNSVQIRNH-LKN 246

RESULT 15
WPA_BACSU
ID WPA_BACSU STANDARD; PRT; 894 AA.
AC P5423; C06726;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cell wall-associated protease precursor (EC 3.4.21.-) [Contains: Cell
GN wall-associated polypeptides CWBP23 and CWBP52].
OS WPA OR BSU10770.
OC Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-54 AND 414-428.
RC STRAIN=168;
RX MEDLINE=97158234; PubMed=9004506;
RA Margot P., Karamata D.;
RT "The wprA gene of Bacillus subtilis 168, expressed during exponential
RL growth encodes a cell-wall-associated protease.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98015415; PubMed=9353931;
RA Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;
RT "Sequencing of regions downstream of addA (98 degrees) and citG (289
RL degrees) in Bacillus subtilis";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Arevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borrius R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Hage K., Haele J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Plescan E., Pujic P., Purnelle B., Rapoport G., Ray M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Srokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Teseo V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RL subtilis";
RN Nature 390:249-256(1997).
CC -!- FUNCTION: NOT YET KNOWN; COULD BE INVOLVED IN PROTEOGLYCAN
CC DEGRADATION, BY CLEAVAGE OF ITS PEPTIDE BRIDGES.
CC -!- SUBCELLULAR LOCATION: Cell-wall bound.
CC -!- PTM: PROCESSED INTO CWBP23 AND CWBP52.

```

Search completed: March 31, 2004, 16:05:33
Job time : 8.49423 secs

```
CC -!- SIMILARITY: Belongs to peptidase family S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U58981; AAC25926.1; -
CC EMBL; Y09476; CAA70641.1; -
CC EMBL; Z99109; CAB12917.1; -
CC FIR; F69730; F69730.
CC HSP; Q45670; IDB1.
CC MEROPS; S08.004; -.
CC Subtilisin; BG11846; wprA.
CC InterPro; IPR000209; Peptidase S8.
CC Pfam; PF00082; Peptidase S8; 1.
CC PRINTS; PR00723; SUBTILISIN
CC PROSITE; PS00136; SUBTILASE ASP; FALSE_NEG.
CC PROSITE; PS00137; SUBTILASE HIS; 1.
CC PROSITE; PS00138; SUBTILASE SER; 1.
CC Hydrolase; Serine protease; Cell wall; Zymogen; Signal;
KW Complete proteome.
FT SIGNAL 1 31
FT CHAIN 32 894 CELL WALL-ASSOCIATED PROTEASE.
FT CHAIN 32 ? CHEP23.
FT PROPEP ? 413 POTENTIAL.
FT CHAIN 414 894 CHEP23.
FT ACT_SITE 466 466 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 497 497 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 650 650 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 9 9 V -> A (IN REF. 1).
FT CONFLICT 14 14 L -> I (IN REF. 1).
FT SEQUENCE 894 AA; 96487 MW; 0F67C353E5F8DBC CRC64;
Query Match 10.8%; Score 243.5; DB 1; Length 894;
Best Local Similarity 23.8%; Pred. No. 1.8e-08;
Matches 110; Conservative 68; Mismatches 150; Indels 135; Gaps 21;
QY 25 IVAVADTGLDTRNDSSMHEAFRGKITALYA--LGRTNNANDPENGHGHVAGSVLGNAT 81
Db 457 LIAVDTGVDTLAD-----LKGKVTLDLGHNFVGRNNANDDQGHGHVAGIITAAQSD 510
QY 82 N-----KGMAPQANLVFQSIMDSCGGLGGLPANLQTLFQAYSAGARIHTNSWGPVNGAY 137
Db 511 NGYSMTGLNAKAKIIPVKVLDSAG--SGDTEQIALGKIYAADKGAKVINLSLG---GGY 564
QY 138 TTDSRVVDYVRK--NDMTILFAAGNEGSGGTISAPGTAKNAITVGATENLRPSFGSYA 195
Db 565 S-----RVLEFALKYAADKNVLIIAAGNDGENALSYPASSKYVMVGAT----- 609
QY 196 DNINHVAQFSRGTRDGRIPKDVMAFGTYILSARS LAPDSF FWHANHDSKYAYMGTSW 255
Db 610 NRDMTADFNSYKGL-----DISAPGSDI-----PSLVNGN-----VTYMSGTSW 651
QY 256 ATPVAGNVAQLREHFVKNRGV--TPKPSLLKALIAAGADYGLGFPNGNQ----- 304
Db 652 ATPYAAAAAGLL---FAQNPKLKRTVEEDMLKKT---ADDISFESVDGGEELDYDYG 704
QY 305 -----GWRVTIDKSLNVAFN--ETSPLSTSQKATYSF----- 336
Db 705 PIETPKTPGVMDHSGYGRNVKMSAADIQLKVNKLESTQTAVRGSAXEGTLIEVMNGK 764
QY 337 -----TAQAGK--PLKISLVMSDAPGSTASLTLVNDLIDLVTAPNGTKYVGNDFTPYDN 390
Db 765 KKLGSAGAKGDNAPKVNIA-----TQKQDQVLYLKATKG----- 798
QY 391 NWDGRNNVNFVINAPOSQTVTVE--VOAYNVVPSPQTFSLAIV 432
Db 799 --DAKTSYKVVVVKSGSGTPKUNAVKTKDTAVKGNKANSKAMI 839
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 32.475 Seconds
(without alignments)

4206.909 Million cell updates/sec

Title: US-09-985-689A-7

Perfect score: 2252

Sequence: 1 NDVARGIVKADVAQNFGLY.....EVQAYNPVSPQTFSLAIVH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea.*
2: SP bacteria.*
3: SP fungi.*
4: SP human.*
5: SP invertebrate.*
6: SP mammal.*
7: SP mhc.*
8: SP organelle.*
9: SP phage.*
10: SP plant.*
11: SP rodent.*
12: SP virus.*
13: SP vertebrate.*
14: SP unclassified.*
15: SP rvirus.*
16: SP bacteriap.*
17: SP archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|--------|---------------------|
| 1 | 2125.5 | 94.4 | 640 | Q93UV9 | Q93UV9 bacillus sp |
| 2 | 2116.5 | 94.0 | 434 | Q9AQR0 | Q9AQR0 bacillus sp |
| 3 | 2110.5 | 93.7 | 639 | Q9AQR3 | Q9AQR3 bacillus sp |
| 4 | 2044 | 90.8 | 433 | Q9AQR1 | Q9AQR1 bacillus sp |
| 5 | 2040 | 90.6 | 433 | Q9AQR4 | Q9AQR4 bacillus sp |
| 6 | 2033 | 90.3 | 433 | Q9AQR2 | Q9AQR2 bacillus sp |
| 7 | 501 | 22.2 | 1825 | Q8T9W1 | Q8T9W1 dictyosteli |
| 8 | 445.5 | 19.8 | 1702 | Q9GTN7 | Q9GTN7 dictyosteli |
| 9 | 404 | 17.9 | 561 | Q8RBJ2 | Q8RBJ2 thermococcus |
| 10 | 396 | 17.6 | 654 | Q8UOC9 | Q8UOC9 pyrococcus |
| 11 | 377 | 16.7 | 430 | Q8ENV1 | Q8ENV1 oceanobacil |
| 12 | 374.5 | 16.6 | 1239 | Q9FBZ4 | Q9FBZ4 streptomyc |
| 13 | 367.5 | 16.3 | 1253 | Q9FC06 | Q9FC06 streptomyc |
| 14 | 355 | 15.8 | 1208 | Q82BI4 | Q82BI4 streptomyc |
| 15 | 341.5 | 15.2 | 1102 | P95684 | P95684 streptomyc |
| 16 | 339.5 | 15.1 | 412 | Q9AER6 | Q9AER6 thermococcus |

| | | | | | | |
|----|-------|------|------|----|--------|--------------------|
| 17 | 339.5 | 15.1 | 412 | 16 | Q8RC68 | Q8RC68 thermoanaer |
| 18 | 334.5 | 14.9 | 444 | 16 | Q9KBJ7 | Q9KBJ7 bacillus ha |
| 19 | 332 | 14.7 | 1139 | 16 | Q82I39 | Q82I39 streptomyc |
| 20 | 329.5 | 14.6 | 1245 | 16 | Q9RL54 | Q9RL54 streptomyc |
| 21 | 327.5 | 14.5 | 1105 | 2 | Q8KKH6 | Q8KKH6 streptomyc |
| 22 | 326.5 | 14.5 | 1220 | 16 | Q9L0A0 | Q9L0A0 streptomyc |
| 23 | 323 | 14.3 | 1237 | 2 | Q8GGT4 | Q8GGT4 streptomyc |
| 24 | 322 | 14.3 | 824 | 2 | Q45464 | Q45464 bacillus sp |
| 25 | 320 | 14.2 | 891 | 1 | O93635 | O93635 thermococcu |
| 26 | 319 | 14.2 | 435 | 16 | Q8EMJ3 | Q8EMJ3 oceanobacil |
| 27 | 316.5 | 14.1 | 442 | 16 | O31788 | O31788 bacillus su |
| 28 | 308 | 13.7 | 1398 | 1 | Q9P9L1 | Q9P9L1 pyrococcus |
| 29 | 293.5 | 13.0 | 1135 | 1 | Q9P9D1 | Q9P9D1 uncultured |
| 30 | 291 | 12.9 | 431 | 2 | Q9S3L6 | Q9S3L6 bacillus sp |
| 31 | 284 | 12.6 | 434 | 2 | O54327 | O54327 bacillus sp |
| 32 | 283.5 | 12.6 | 799 | 16 | Q9KEM1 | Q9KEM1 bacillus ha |
| 33 | 282 | 12.5 | 959 | 16 | Q9PMS7 | Q9PMS7 xanthomonas |
| 34 | 277.5 | 12.3 | 1345 | 1 | O54437 | O54437 staphylothe |
| 35 | 275.5 | 12.2 | 1571 | 2 | Q8GCW3 | Q8GCW3 streptococ |
| 36 | 275 | 12.2 | 814 | 16 | Q82VB3 | Q82VB3 nitrosomona |
| 37 | 274.5 | 12.2 | 1098 | 16 | Q9L128 | Q9L128 streptomyc |
| 38 | 268.5 | 11.9 | 1570 | 16 | Q8E2V6 | Q8E2V6 streptococ |
| 39 | 268.5 | 11.9 | 1570 | 16 | Q8DX06 | Q8DX06 streptococ |
| 40 | 266 | 11.8 | 966 | 16 | Q8PB28 | Q8PB28 xanthomonas |
| 41 | 265 | 11.8 | 403 | 2 | Q45463 | Q45463 bacillus sp |
| 42 | 265 | 11.8 | 419 | 2 | Q45681 | Q45681 bacillus su |
| 43 | 264.5 | 11.7 | 586 | 16 | Q8PAL8 | Q8PAL8 xanthomonas |
| 44 | 262.5 | 11.7 | 715 | 2 | P70765 | P70765 alteromonas |
| 45 | 261.5 | 11.6 | 627 | 16 | Q9RUD0 | Q9RUD0 deinococcus |

ALIGNMENTS

RESULT 1

Q93UV9 PRELIMINARY; PRT; 640 AA.
AC Q93UV9;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Protease.
GN PROF.
OS Bacillus sp. KSM-KP43.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=109322;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KP43;
RA Itoh S., Saeki K.;
RT "new protease."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB051423; BAB55674.2; -
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; P:peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
SQ SEQUENCE 640 AA; 67991 MW; 48BAF77E9D592C15 CRC64;

Query Match 94.4%; Score 2125.5; DB 2; Length 640;
Best Local Similarity 93.5%; Pred. No. 3.3e-116;
Matches 406; Conservative 19; Mismatches 8; Indels 1; Gaps 1;

Qy 1 NDVARGIVKADVAQNFGLYGQGIIVAVADTGLDTCGRNDSMHEAFRGKITALLYALGRTN 60

Db 207 NDVARGIVKADVAQNFGLYGQGIIVAVADTGLDTCGRNDSMHEAFRGKITALLYALGRTN 266

Qy 61 NANDPNGHGHVAGSVLGN-ATNKMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 119

```
Db 267 NNDTNGHGTAVAGSVLNGSTNKGMAPOANLVFQSIMDSGGGLGLPSNLQTLFSQAYS 326
QY 120 AGARHTNSWGPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179
Db 327 AGARHTNSWGAANGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 386
QY 180 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPTIYLSARSSLPDSSF 239
Db 387 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPTIYLSARSSLPDSSF 446
QY 240 WANHDSKYAYMGSTMATPIVAGNVAQIREHFVNKGVTPEKPSLLKAALTAGADVGLGF 299
Db 447 WANHDSKYAYMGSTMATPIVAGNVAQIREHFVNKGVTPEKPSLLKAALTAGADVGLGF 506
QY 300 PNGNQGWGRVTLDKSLNVAFYNETSPLSTSQATYSFTAQAGKPLKISLVMSDAPGSTTA 359
Db 507 PNGNQGWGRVTLDKSLNVAFYNETSPLSTSQATYSFTAQAGKPLKISLVMSDAPGSTTA 565
QY 360 SLTLVNDLVLITAPNGTKYVNDFTAPYDNNWDRNNVNFVINAPOSQTYTVEVOAYN 419
Db 567 SVTLVNDLVLITAPNGTKYVNDFTAPYDNNWDRNNVNFVINAPOSQTYTVEVOAYN 626
QY 420 VPVSPQTFSLAIVH 433
Db 627 VPVSPQTFSLAIVN 640

RESULT 2
Q9AQR0 PRELIMINARY; PRT; 434 AA.
AC Q9AQR0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protease (Fragment).
GN PROE.
OS Bacillus sp. NV1.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=133781;
RN [1]_TaxID=133781;
SEQUENCE FROM N.A.
RC STRAIN=NV1;
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus sp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319 (2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR EMBL; AB046406; BAB21269.1; -.
DR HSSP; P00782; 1SUP.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002029; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1
FT TER 434
SQ SEQUENCE 434 AA; 45294 MW; 83517EDDB74125D2 CRC64;
```

```
Query Match 94.0%; Score 2116.5; DB 2; Length 434;
Best Local Similarity 92.2%; Pred. No. 6.7e-116;
Matches 400; Conservative 21; Mismatches 12; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVQNNGFLYGGQGVAVADTGLDTRNDSSMHEAPRGKITAYALGRTN 60
```

```
Db 1 NDVARGIVKADVQSSYGLYGGQGVAVADTGLDTRNDSSMHEAPRGKITAYALGRTN 60
QY 61 NANDPNGHGTAVAGSVLNGAT-NKGMAPOANLVFQSIMDSGGGLGLPANLQTLFSQAYS 119
Db 61 NANDPNGHGTAVAGSVLNGAT-NKGMAPOANLVFQSIMDSGGGLGLPSNVSTLFSQAYS 120
QY 120 AGARHTNSWGPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179
Db 121 AGARHTNSWGPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 180
QY 180 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPTIYLSARSSLPDSSF 239
Db 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPTIYLSARSSLPDSSF 240
QY 240 WANHDSKYAYMGSTMATPIVAGNVAQIREHFVNKGVTPEKPSLLKAALTAGADVGLGF 299
Db 241 WANHDSKYAYMGSTMATPIVAGNVAQIREHFVNKGVTPEKPSLLKAALTAGADVGLGF 300
QY 300 PNGNQGWGRVTLDKSLNVAFYNETSPLSTSQATYSFTAQAGKPLKISLVMSDAPGSTTA 359
Db 301 PNGNQGWGRVTLDKSLNVAFYNETSPLSTSQATYSFTAQAGKPLKISLVMSDAPGSTTA 360
QY 360 SLTLVNDLVLITAPNGTKYVNDFTAPYDNNWDRNNVNFVINAPOSQTYTVEVOAYN 419
Db 361 SVTLVNDLVLITAPNGTKYVNDFTAPYDNNWDRNNVNFVINAPOSQTYTVEVOAYN 420
QY 420 VPVSPQTFSLAIVH 433
Db 421 VPOGPOAFSLAIVN 434

RESULT 3
Q9AQR3 PRELIMINARY; PRT; 639 AA.
AC Q9AQR3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protease.
GN PROA.
OS Bacillus sp. 9860.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=133778;
RN [1]_TaxID=133778;
SEQUENCE FROM N.A.
RC STRAIN=9860;
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus sp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319 (2000).
DR EMBL; AB046403; BAB21266.2; -.
DR HSSP; P00782; 1SUP.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002029; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
SQ SEQUENCE 639 AA; 68185 MW; 316AF6PDBE4FF54 CRC64;
```

```
Query Match 93.7%; Score 2110.5; DB 2; Length 639;
Best Local Similarity 93.1%; Pred. No. 2.5e-115;
Matches 404; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVQNNGFLYGGQGVAVADTGLDTRNDSSMHEAPRGKITAYALGRTN 60
```

Db 206 NDVARGIVKADVQAQSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRIN 265
 QY 61 NNDPNGHGHGTHVAGSVLGN-ATNKMAPQANLVFQSIMDSGGGLGGLPANLQTLFQAYS 119
 Db 266 NNDPNGHGHGTHVAGSVLGN-ATNKMAPQANLVFQSIMDSGGGLGGLPANLQTLFQAYS 325
 QY 120 AGARIHNSGAPVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179
 Db 326 AGARIHNSGAAVNGAYTTDSRNVDDYVRKNDMTILFAAGNERPNSGTISAPGTAKNAI 385
 QY 180 TVGATENLRPSFGSVADNINHVAFSSRPGTDRGRIKPDVMACTYILSAPSSILAPDSSF 239
 Db 386 TVGATENLRPSFGSVADNINHVAFSSRPGTDRGRIKPDVMACTYILSAPSSILAPDSSF 445
 QY 240 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAAGADVGLGF 299
 Db 446 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAAGADVGLGF 505
 QY 300 PNGNGGWRVTLDKSLNAVFNETSPLSTSKATYSFTAQAGKPLKISLVMSDAPGTTA 359
 Db 506 PNGNGGWRVTLDKSLNAVFNETSPLSTSKATYSFTAQAGKPLKISLVMSDAPGTTA 565
 QY 360 SLTVNDLDELVTAPNGTKYVGNDFTPYDNNWDRNNVNFVFNAPSGTYTVEVQAYN 419
 Db 566 SLTVNDLDELVTAPNGTKYVGNDFTPYDNNWDRNNVNFVFNAPSGTYTVEVQAYN 625
 QY 420 PVSPTQFSLAIVH 433
 Db 626 PVSPTQFSLAIVH 639

RESULT 4

Q9AQR1 PRELIMINARY; PRT; 433 AA.
 AC Q9AQR1; (TREMELrel. 17, Created)
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Protease (Fragment).
 GN PROD.
 OS Bacillus sp. SD521.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=133780;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SD521;
 RX MEDLINE=20568675; PubMed=11118284;
 RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
 Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from
 RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
 RT evolutionary relationships.";
 RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 DR EMBL; AB046405; BAB21268.1; -
 DR HSSP; Q45670; IDBI.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydrolyase; Protease; Serine Protease.
 FT NON_TER 1
 FT NON_TER 433
 SQ SEQUENCE 433 AA; 45576 MW; 98A2DF18FE66DDC CRC64;

Query Match 90.8%; Score 2044; DB 2; Length 433;
 Best Local Similarity 89.1%; Pred. No. 1.1e-111;

Matches 386; Conservative 24; Mismatches 23; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVQAQSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRIN 60
 Db 1 NDVARGIVKADVQAQSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRIN 60
 QY 61 NNDPNGHGHGTHVAGSVLGN-ATNKMAPQANLVFQSIMDSGGGLGGLPANLQTLFQAYS 120
 Db 61 NNDPNGHGHGTHVAGSVLGN-ATNKMAPQANLVFQSIMDSGGGLGGLPANLQTLFQAYS 120
 QY 121 GARIHNSGAPVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 180
 Db 121 GARIHNSGAPVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 180
 QY 181 VGATENLRPSFGSVADNINHVAFSSRPGTDRGRIKPDVMACTYILSAPSSILAPDSSF 240
 Db 181 VGATENLRPSFGSVADNINHVAFSSRPGTDRGRIKPDVMACTYILSAPSSILAPDSSF 240
 QY 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAAGADVGLGF 300
 Db 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAAGADVGLGF 300
 QY 301 PNGNGGWRVTLDKSLNAVFNETSPLSTSKATYSFTAQAGKPLKISLVMSDAPGTTA 360
 Db 301 PNGNGGWRVTLDKSLNAVFNETSPLSTSKATYSFTAQAGKPLKISLVMSDAPGTTA 360
 QY 361 SLTVNDLDELVTAPNGTKYVGNDFTPYDNNWDRNNVNFVFNAPSGTYTVEVQAYN 420
 Db 361 SLTVNDLDELVTAPNGTKYVGNDFTPYDNNWDRNNVNFVFNAPSGTYTVEVQAYN 420
 QY 421 PVSPTQFSLAIVH 433
 Db 421 PVSPTQFSLAIVH 433

RESULT 5
 Q9AQR4 PRELIMINARY; PRT; 433 AA.
 AC Q9AQR4; (TREMELrel. 17, Created)
 DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Protease (Fragment).
 GN PROD.
 OS Bacillus sp. D6.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=127889;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D6;
 RX MEDLINE=20568675; PubMed=11118284;
 RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
 Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from
 RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
 RT evolutionary relationships.";
 RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 DR EMBL; AB046402; BAB21265.1; -
 DR HSSP; Q45670; IDBI.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydrolyase; Protease; Serine protease.
 FT NON_TER 1
 FT NON_TER 433

```

SQ SEQUENCE 433 AA; 45636 MW; 52087E0A2516107P CRC64;
Query Match 90.6%; Score 2040; DB 2; Length 433;
Best Local Similarity 88.9%; Pred. No. 2e-111;
Matches 385; Conservative 24; Mismatches 24; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQNNGFLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRN 60
DB 1 NDVARGIVKADVAQNNGFLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRN 60
QY 61 NANDPNGHGTTHVAGSVLGNATKNGAPQANLVFQSIMDSSGGGLGLPANLQTLFSQAYSA 120
DB 61 NANDPNGHGTTHVAGSVLGNATKNGAPQANLVFQSIMDSSGGGLGLPANLQTLFSQAYSA 120
QY 121 GARIHTNSGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
DB 121 GARIHTNSGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
QY 181 VGATENLRPSFGSYADNINHVAQFSSRGTRGRKIPDVAPGTIVLSARSSLAPDSSFW 240
DB 181 VGATENLRPSFGSYADNINHVAQFSSRGTRGRKIPDVAPGTIVLSARSSLAPDSSFW 240
QY 241 ANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAGAADVGLGFP 300
DB 241 ANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAGAADVGLGFP 300
QY 301 NGNQGRVTLDKSLNVAFYNETSPSTSKATYSFTAQAGKPLKISLVMSDAPGTTAS 360
DB 301 NGNQGRVTLDKSLNVAFYNETSPSTSKATYSFTAQAGKPLKISLVMSDAPGTTAS 360
QY 361 LTVLNDLVLITAPNGTKVGNDFTPYDNNNDGRNNVFNAPQSGTYTVEQAYNV 420
DB 361 LTVLNDLVLITAPNGTKVGNDFTPYDNNNDGRNNVFNAPQSGTYTVEQAYNV 420
QY 421 PVSQPTFSLAIVH 433
DB 421 PVSQPTFSLAIVH 433
RESULT 6
Q9AQR2 PRELIMINARY; PRT; 433 AA.
AC Q9AQR2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protease (Fragment).
GN PROC.
OS Bacillus sp. Y.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_TaxID=133779;
RN [1]
RP SEQUENCE FROM N.A.
PC STRAIN=Y.
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K.; Okuda M.; Hatada Y.; Kobayashi T.; Ito S.; Takami H.;
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships."
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR EMBL; AB046404; BAB21267.1; -.
DR HSP; Q45670; IDBI.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptide_S8; 1.
DR Pfam; PF04151; PFC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1 1
FT NON_TER 433 433
SQ SEQUENCE 433 AA; 45587 MW; B81291A803C775AE CRC64;
Query Match 90.3%; Score 2033; DB 2; Length 433;
Best Local Similarity 88.7%; Pred. No. 5e-111;
Matches 384; Conservative 25; Mismatches 24; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQNNGFLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRN 60
DB 1 NDVARGIVKADVAQNNGFLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRN 60
QY 61 NANDPNGHGTTHVAGSVLGNATKNGAPQANLVFQSIMDSSGGGLGLPANLQTLFSQAYSA 120
DB 61 NANDPNGHGTTHVAGSVLGNATKNGAPQANLVFQSIMDSSGGGLGLPANLQTLFSQAYSA 120
QY 121 GARIHTNSGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
DB 121 GARIHTNSGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
QY 181 VGATENLRPSFGSYADNINHVAQFSSRGTRGRKIPDVAPGTIVLSARSSLAPDSSFW 240
DB 181 VGATENLRPSFGSYADNINHVAQFSSRGTRGRKIPDVAPGTIVLSARSSLAPDSSFW 240
QY 241 ANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAGAADVGLGFP 300
DB 241 ANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAGAADVGLGFP 300
QY 301 NGNQGRVTLDKSLNVAFYNETSPSTSKATYSFTAQAGKPLKISLVMSDAPGTTAS 360
DB 301 NGNQGRVTLDKSLNVAFYNETSPSTSKATYSFTAQAGKPLKISLVMSDAPGTTAS 360
QY 361 LTVLNDLVLITAPNGTKVGNDFTPYDNNNDGRNNVFNAPQSGTYTVEQAYNV 420
DB 361 LTVLNDLVLITAPNGTKVGNDFTPYDNNNDGRNNVFNAPQSGTYTVEQAYNV 420
QY 421 PVSQPTFSLAIVH 433
DB 421 PVSQPTFSLAIVH 433
RESULT 7
Q8T9W1 PRELIMINARY; PRT; 1825 AA.
AC Q8T9W1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serine protease/ABC transporter Tagd.
GN TAGD.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
CX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
PC STRAIN=Ax4;
RA Anjard C.; Loomis W.F.;
RT "Evolution of the ABC transporters of Dictyostelium."
RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; AF46309; AAL74253.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0001666; F:nucleotide binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_Arpase.
DR InterPro; IPR001140; ABC_TM_transpt.
```

```

DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR00209; Peptidase S8.
DR Pfam; PF00664; ABC membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PRODOM; PD00006; ABC transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW ATP-binding; Protease; Transport.
SQ SEQUENCE 1825 AA; 202641 MW; E281608C78613A3B CRC64;

Query Match      22.2%; Score 501; DB 5; Length 1825;
Best Local Similarity 27.9%; Pred. No. 1.1e-20;
Matches 169; Conservative 73; Mismatches 155; Indels 208; Gaps 23;

QY 11 DVAQNN-----FGLYQGGQIVAVDTGLDGTGR---NDS-----SMHEAFRGKITALYAL 56
DB 315 DTLVNNDRIDIPLRGKGQILSTADTGLDGHCHFFSDSNPIPNVSNVNLNHRKVYTYIGSL 374
QY 57 GRTNANDPNHGHTHVGASVLGN-----ATNKGMAPQANLVFQSI-MDSGGGLGGLP 107
DB 375 --HDNEDVVDHGHTHVGCSAGAPEDDSLAISFSGLATDAKIAFFDLADPSNNPEVPP 432
QY 108 ANLQTLFQSAQSAGARIHTNSGWA-----PVNGAYTDSRNVDYVRKN-DMTILFAAGNE 162
DB 493 EDSYQLYPLYNAGARVHGDSWGLSIQYGLGSYDDAGSIDDEFLYHPDPIILRAAGNN 492
QY 163 GPGSGTISAPGTAKNAITVGATENLRPSF-----
DB 493 EGYSSLLS-QATAKNVITVGAETHTHSTYTDALREYSNPFVAKSTLNSLCQSFDDKYCT 551
QY 192 -----GSYAD-----NINHAQFSRRGFTDRGRKP 217
DB 552 YTTAQCTEYSTVKGLSGCCTSYIKNSYASIFSSQPELYNENNICSFSKGPETHDGRKP 611
QY 218 DYMAEGTYLSARSLSA-----PDSFWANHDSKYAYMGTSMATPIVAGNVAQ 266
DB 612 DIVAPQVITTSARSNGANTDOCGDGLPNTNALLSE-----SGTSMATPLATAATTI 664
QY 267 LREH-----FVAKRGVTPKPSLLKAALIAAGAADVGLGFP----- 300
DB 665 LRQYLDGVGYPTGSIYESNKLQPTGSLKALMINNAQLNGTFPLSSNTNPSNAVPTDF 724
QY 301 ----NGNQHGRTLDKSLNVAFVNETS-----PLTSQ 330
DB 725 AGANFVQGWGLRMSEWLYV----BSSGVKPKPSRWVGIGELGDKKASNNWKEYSLSTGQ 780
QY 331 KATYGTFAQ-----AGKP-LKISLVNSDAPGTTASLTLVNLDLVLIT----- 372
DB 781 NVSYCTFYKPSSSGNSGGIPIVATLVMTDPPSYSGAKLVNLDLTWNTSESEIFY 840
QY 373 -----APNGTKYVGNDFAPYDNNWGRNNVNVF-----INAPQGTYYVEQYVNVPS 423
DB 841 SNSGSSSYNGTK-----GTTLPQ-----DSINNVEGIITYPTNKSEISFRFIAGTNPIG 893
QY 424 PQTFSS 428
DB 894 PQNFS 898

RESULT 8
Q9GTN7 PRELIMINARY; PRT; 1702 AA.
AC Q9GTN7
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Taga.
GN TAGA.

```

```

OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]_SEQUENCE FROM N.A.
RP Good J.R., Cabral M., Kuspa A.;
RA "Taga, a putative serine protease/ABC transporter of Dictyostelium
RT that is expressed at the onset of development and is required for the
RT differentiation of a subpopulation of prespore cells.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; AF263455; AAG11416.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABC_TM_transpt.
DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR000209; Peptidase S8.
DR Pfam; PF00664; ABC membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PRODOM; PD00006; ABC_TRANSPORTER; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Transport.
SQ SEQUENCE 1702 AA; 187103 MW; 4A67716303CB7131 CRC64;

Query Match      19.8%; Score 445.5; DB 5; Length 1702;
Best Local Similarity 26.2%; Pred. No. 1.7e-17;
Matches 142; Conservative 97; Mismatches 163; Indels 141; Gaps 23;

QY 18 GLYQGGQIVAVADTGLDGTGR---NDSMHEAFRGKITALYALCRTNANDPNHGHTHVG 74
DB 300 GIKGDEIVGCADTGIDINHCFYDTNPIGSTHRKIIS-YSSGNGDQIDEIGHGTHVG 358
QY 75 SVLGNAT-----NKGMAPQANLVFQSI-MDSGGGLGGLPANLQTLFQSAQSAGARIHT 126
DB 359 TIIGSTTVDPVSSEFSGGAPNSKAVFDVLQVSGNGL-SIQSNLTAYISTYDQNAKVHC 417
QY 127 NSNGAPVNGAYTTDSRNVDYVRKN-DMTILFAAGNE---GPGS-GTISAPGTAKNAITV 181
DB 418 DAWNMGIFGYTGVTEMIDRFQWDHDFLVVRSAGNNVFNFSIYTLQSSETSKNSLV 477
QY 182 GATENLRPSFGSYADNINH----- 200
DB 478 GSSN--QPS-STYLSIDYMDWDFYNSIRTSVCTQGSIYGITCSDVPTQTSVDIQ 534
QY 201 -----VAQFSRRGPTDRGRKPKDPMAGPTVLSA 229
DB 535 CCNPNILAKICCTEIQOQYQNTSVYSEFISLPSGVGPTSDGRKLPDLPAGPSIIS 594
QY 230 RSSLAPDSSFWANH-----DSKYAYMGTSMATPIVAGNVAQIRE-----HFVKNR 275
DB 595 R-SLGPSSSTI--NHCSPITSGIATSAIAMESSQAAVAATSAVLVQRYDGYFNGK 651
QY 276 -----GVTPKPSLLKAALTAGAA---DVLGFPNGQHGRTLDKSLNVAFVNETS--- 324
DB 652 VNSGVQFQPSASLVKATLNTASINVDSTLEY---SQGFNIQLSKLITTTNAQTSLDI 708
QY 325 -----PLTSQKATYSFTAAQKPLKISLVNSDAPGTTASLTLVNLDLVLIT-P 374
DB 709 PSSIEKADPIINTGETNSYCFSLDSKADIDITLVMTDPPAGSPLSTFLVNNLDLALLAF 768
QY 375 NG--TKYVGNDFAPYDNN-----WDGRNNVNVFINAPQGTYYVEQYVNVPSQTFSS 428
DB 769 DGBLSIYSGNSETI-PKNTSQVIFDQNLNNVEIRINDAPIGSYDVVKIFGTNIVIPNQSYS 827

```

```

Qy 429 LAI 431
Db 828 VVI 830

RESULT 9
Q8RBJ2 PRELIMINARY; PRT; 561 AA.
ID Q8RBJ2
AC Q8RBJ2
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Subtilisin-like serine proteases.
GN APR22 OR TTE0824.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=2192816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; A5013049; AAM24081.1;
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR002029; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR InterPro; IPR009020; Protease_inhib.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Protease; Complete proteome.
SQ SEQUENCE 561 AA; 59968 MW; BA9C5C52F7083A18 CRC64;

Query Match 17.9%; Score 404; DB 16; Length 561;
Best Local Similarity 30.2%; Pred. No. 1.1e-15;
Matches 133; Conservative 65; Mismatches 162; Indels 80; Gaps 16;

Qy 6 GIVKADVAQNNGLYGQGIIVAVADTGLTGRNDSMHEAFRGKITALYALGRTNAN-- 63
Db 155 GITK--ASRDFGVTGKNTIATIDTGDGNHVDLS-----GKGI-----IGWKDFINNK 201

Qy 64 ----DPNGHGHV----AGSVLGNATNKGMAPQANLVFQSIM--DSGGGLGLPANLQTL 113
Db 202 TTPYDDNGHGHVASTAAGTGAGNSFYKGVAFDALLVIGIKVLADANGSGSMSTVTAGIDWA 261

Qy 114 PSQAYSAGARIHNSGAPVNGAYTDSRNVDDYVRKNDMTILFAAGNEGPGSGTISAPG 173
Db 262 VQNKQVYGVKVINLSLGTSTSDGTDSTSLAVNRAVDSGIVVVVAAGNSGPAKYTIGSPG 321

Qy 174 TAKNAITVGTATENLRPSFSGYADNINHVAFQSSRGPTDRGRKPDVMAFGTYILSARSSL 233
Db 322 AAEKAITVAAMADV----GELGFNL--ASFSSRGPTADGRKPKDIAAPGYNITAAK--- 371

Qy 234 APDSSFWANHDSKYAMGTSMTATPIVAGNVLAQLREHFVNGKVTGTPKPSLLKAALIA 293
Db 372 -----ANSVNGYVTVSGTSMTATPEVAGTVLMLN---ANENLTPNDA--KNIIINSTAK 419

Qy 294 DVGLGFPNGNOGWRVTLDKSLNVA-----FVNETSPLTSQKATYSFTA- 338

```

```

Db 420 SWGPPSKNVYDYGAGRLDGYEALRVAGNFRGNNDIVPNHYVYISGYLP--GSRYSDTWTENAT 478
Qy 339 QAGKPKLSLVNSDAPGSTTASLTILVNDLDLVITAPNGTKYVGNDFTPADYNNWGRNRV 398
Db 479 NTSYPIAITLIIPDWANTNP-----DFDIYLYDPSGLTIKSS-----TGTORQ 521

Qy 399 ENVFINAPQSQTVEVOAY 418
Db 522 EITILPSQTGYVYKVS 541

RESULT 10
Q8UOC9 PRELIMINARY; PRT; 654 AA.
ID Q8UOC9
AC Q8UOC9
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Alkaline serine protease.
GN PF1670.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb P.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; A6010265; AAL81794.1;
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002029; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Protease; Complete proteome.
SQ SEQUENCE 654 AA; 70230 MW; 1CB145A5F505DB34 CRC64;

Query Match 17.6%; Score 396; DB 17; Length 654;
Best Local Similarity 29.5%; Pred. No. 3.8e-15;
Matches 127; Conservative 56; Mismatches 163; Indels 84; Gaps 15;

Qy 16 NFGLYQGGQIVAVADTGLTGRNDSMHEAFRGKITALYALGRTNAN-----DPNGHG 69
Db 152 NLGYDGGITIGITIDTGD-----ASHPDLQGV-----IGWDFVNGRSYPYDDHGHG 200

Qy 70 THVAGSVLG-----NATNKGMAPQANLVFQSIM--DSGGGLGLPANLQTLFSQAYSAGA 122
Db 201 THVASTAAGTGAASNGKYKGMAPQAKLIGIKVLGADGSGSISTIKGVFWAVNDKDKYGI 260

Qy 123 RIHTNSGAPVNGAYTDSRNVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAITVG 182
Db 261 KVINLSLGSQSSDGTDALSQAVNAADAGLVVVVAAGNSGPNKYTIGSPAAASKVITVG 320

Qy 183 ATENLRPSFSGYADNINHVAFQSSRGPTDRGRKPDVMAFGTYILSARSLAPDSSFWAN 242
Db 321 A-----VDKYDVTITFSRSGPTADGRUKPEVAPAGNWIIAARAS----GTSMGQP 366

Qy 243 HDSKYATMGTSMTATPIVAGNVLAQLREHFVNGKVTGTPKPSLLKAALIAAG----- 292
Db 367 INDYTTAAPTGSMTATPHVAGIAALLQ-----AHPSTPDKVKTALISTADIVKPEI 419

Qy 293 ADVGLGFPNGNOGWRVTLDKSL---NVAFVNETSPLTSQKATYSFTAQAGKPLKISLV 349
Db 420 ADIAYGA-----GRVNAKATNYAKLVFTGYVANKGSGTHQFVIGSGASFTATLY 472

Qy 350 WSDAPGSGTTASLTILVNDLDLVITAPNGTKYVGNDFTPADYNNWGRNRVFNAPQSG 409

```

```
Db 473 WDAN-----SDLDLYLDNGNQ-VDYSYTAAY-----GFEKVGYNPTDG 513
QY 410 TYTVEYQAYN 419
Db 514 TWTIKVVSYS 523

RESULT 11
Q8ENV1 PRELIMINARY; PRT; 430 AA.
AC Q8ENV1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Intracellular alkaline serine proteinase.
GN OB2375.
OS Oceanobacillus iheyensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX MEDLINE=22220767; PubMed=12235376;
RA Takami H., Takaki Y., Uchiyama I.;
RT Ridge and its unexpected adaptive capabilities isolated from the Iheya
RT environments."
RT Nucleic Acids Res. 30:3927-3935 (2002).
DR EMBL; AP004601; BAC14331.1; -
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Complete proteome.
SQ SEQUENCE 430 AA; 45838 MW; 6D09A99BBC1E310F CRC64;

Query Match 16.7%; Score 377; DB 16; Length 430;
Best Local Similarity 34.7%; Pred. No. 2.8e-14;
Matches 118; Conservative 44; Mismatches 122; Indels 56; Gaps 15;

QY 2 DVARGIVKADVACNPGFLYGQGVAVADTGLDTRNDSSMEAFRGKIT--ALVALGET 59
Db 121 DTASSINADVLES-GLTGQSGSTIAVDITGIHP-----HEDLEGRIGFADFVKGT 172
QY 60 NNANDPNHGHTHVAGSVLGNAT-----NKGMAPOANLVFQSTMDSGGIGGLPA----- 108
Db 173 EPYDD-NGHGTGHCAGDAAGNALSDGQYQGPAPDANLVGVKVLNKTGS-GSLSTVIEGID 230
QY 109 -----NLQTLFQAYSAGARIHTNSGAPVNGAYTTDSRNVDDVYRKNDMTILFAA 159
Db 231 WCIONQSKYNINIL-----SLSGSDATEPAGDFPVNAVET-----AWDNGVVCVAA 279
QY 160 GNEGPGSGTISAPGTAKNAITVGATENLRPSGVSADNINHVAFSSRGPTRDGRIKPDV 219
Db 280 GNSGPGDKTVGSPCLSPKVTIVGADDDNTAERS-----DDVAEFSRSGPTIDGLTKNL 335
QY 220 MAPGTIVLSARS--SLAPSSFWANHSKYAYVGTSMATPIVAGVAQLREHFVNKGV 277
Db 336 LTPGVDIVSLRPGSPIDTKNSARVGSNYISLSGTSMATPICAGVAILQ---SDSSL 392
QY 278 TPKEPSLLKALAGADVGLGFENGQNGRVTLDKSLNV 317
Db 393 T--PNQVKEKLMACQDLQGS-FN-VQAGLYNANLINI 428

RESULT 12
Q9FBZ4 PRELIMINARY; PRT; 1239 AA.
AC Q9FBZ4;
```

```
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative secreted peptidase.
GN SC07188 OR SC8A11.16C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Saunders D.C., Harris D.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.;
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RT Mol. Microbiol. 21:177-96 (1996).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147 (2002).
RL EMBL; AL939130; CAC01588.1; -
DR HSRF; Q99405; IMP1.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR003137; PA.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF02225; PA; 1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Complete proteome.
SQ SEQUENCE 1239 AA; 128505 MW; 8F5E9AC68E31260A CRC64;

Query Match 16.6%; Score 374.5; DB 16; Length 1239;
Best Local Similarity 29.2%; Pred. No. 1.6e-13;
Matches 137; Conservative 57; Mismatches 180; Indels 95; Gaps 15;

QY 8 VKADVACNPF-----GLYGQGVAVADTGLDTRNDSSMEAFRGKITAYALG 57
Db 219 VEADLADSTAGIAPRAWAGNTGGQVEVAVLDTGVDAG-----HPDLADRTAARQSFV 272
QY 58 RTNNANDPNHGHTHVAGSVLGNAT-----NKGMAPOANLVFQSTMDSGGIGGLPANIQT 112
Db 273 PDENTDDRDGHGTHTVASTIAGTGAASAGKEGVAPGARLSIGKVLDN-SRGGQISWTAA 331
QY 113 LFSQAYSAGARIHTNSGA-----PVNGAYTTDSRNVDDVYRKNDMTILFAAGNEGP 164
```

Db 332 MEWAVERHAKIVNLSGEGSDSDPM-----SRVDRLSAQTCALFVVAAGNGGE 384
 QY 165 SSGTISAPCTAKNAITVGATENLRPSFGSYADNINHVAQFSRGPTRDGRKPDVMAFGT 224
 Db 385 -AGSIGAPGVATSALTGVA-----VDATDTLAPFSSQGPVRDGLAKPEITAPGV 432
 QY 225 YILSARSLAPDSSFWANHDSKYAYMGTSNATPIVAGNVAQLREHFVKNRGVTPKPSLL 284
 Db 433 GILAA-----NSSFAAGNGAYQSLSGTSNATPHVAGAAALL-----AAAPDLS 477
 QY 285 KAAL---IAGAADVGLGPNQNGRVRTLDKSLN-----VAFVNETSPSTSQKATY 334
 Db 478 GSAKADVLASSHRTPRYDAFQAGSGRVDVDAVRAGVYASATAYAPGSSFGPVRLVTV 537
 QY 335 STFAQAGKPLKISLWSDAPGT-----TASLTUNLDLVLITAPNGTKYV 380
 Db 538 TTTTGAAVTLELSVAATHAPEGVFLKSASRVTVPAHGTAADVTLTIDGS---GSAGGRAYS 594
 QY 381 GNDFTAPYDNNWGRNVNVFINAPQSGTYVEQVAYNVPSQTFSL 429
 Db 595 GQILA-----TDADARNVAHTAVSAGFVRHKLTVHFKDADGNPVGPFVFL 639

RESULT 13

Q9FC06
 ID Q9FC06 PRELIMINARY; PRT; 1253 AA.
 AC Q9FC06;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Putative secreted peptidase.
 GN SC07176 OR SC8A11.04C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Saunders D.C., Harris D.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RA "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RA MEDLINE=21996410; PubMed=8843436;
 RX Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RA "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)."
 RL Nature 417:141-147(2002).
 RL EMBL; AL939130; CAC01576.1; -.

DR HSSP, Q99405; IMPT.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR003137; PA.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF02225; PA; 1.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PRO0723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Complete proteome.
 SQ SEQUENCE 1253 AA; 130971 MW; AA69B417EFEDB89 CRC64;

Query Match 16.3%; Score 367.5; DB 16; Length 1253;
 Best Local Similarity 28.9%; Pred. No. 4.1e-13;
 Matches 140; Conservative 60; Mismatches 171; Indels 113; Gaps 22;

QY 18 GLYGGQGIIVAVADTGLDGRNDSMHEAFRGKITALYALGRNTNANDPNHGHTHVAGSYL 77
 Db 235 GNTGEGGVAVLDTGVDAG-----HPDFAGRIATAATASFPDQVTDNRNGHGHVASTVA 288
 QY 78 G-----NATNKGMAPOANLVFQSIMDSGG-----LGGLPANLQTLFSQAYSAGARIHT 126
 Db 289 GTGAASGGVEKGVAPGASLHGKVLNDSGSGDSWVLAMWAVR-----DOHAKIVS 341
 QY 127 NSWG-AFVNGAYTTD--SRNVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAITVGA 183
 Db 342 MSLGDSPTDG---TDPLSEAVNWLSAETGALFVVAAGNSGPEAYTVGTTPAAADAALTVGA 398
 QY 184 TENLRPSFGSYADNINHVAQFSRGP-TRDGRKPDVMAFGTYILSARSLAPDSSFWAN 242
 Db 399 VNG--FGKG-----VDQLADFSSRGPVRGDNVAKPDLTAPGVGLAARSRYAPEG----- 446
 QY 243 HDSKYAVMGTSNATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALAGAADVGLGFPNG 302
 Db 447 -EGAYQSLSGTSNATPHVAGAAALLAAEHPDWTG-----QRLKEALVGTAGT-QRFSFP 499
 QY 303 NQNGRVRTLDKSLN-----AFVNETSPSTSQ--KATYFTAQAGKPLKISLWSDA 353
 Db 500 DAGSGRVDVAAAVRSTILASGDFAQAHVPTPGQTVRRDVTYNSGPAFVALDLALSPA 559
 QY 354 -----PGSTTASLTUNLDL-----LVITAPNGTKVGNDFTA 386
 Db 560 ELPEGLFTLSEAQVTPAHGTASGVITHLDAEDNGAYATRLVASGADGAVLA----RT 615
 QY 387 PYDNNWGRNVNVFINA-----FQSGT-----YTVEVQA-VNVPVSPQ 425
 Db 616 PVGVNKEGRR--ATLALTAKDHHDKPLSGTVILKDVERTNAPKVYSVDASGRLLDLRLSPS 673
 QY 426 TFSL 429
 Db 674 TYSV 677

RESULT 14

Q82B14
 ID Q82B14 PRELIMINARY; PRT; 1208 AA.
 AC Q82B14;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative peptidase.
 GN SAV5721.
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948;

RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: deducing the ability of producing secondary
 RT metabolites.",
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=12692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.,
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermitilis.",
 RL Nat. Biotechnol. 21:526-531(2003).
 DR EMBL; AF005044; BAC73433.1;
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR003137; PA.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF02225; PA; 1.
 DR PRINTS; PR00082; Peptidase S8; 1.
 DR PROSITE; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Complete proteome.
 SQ SEQUENCE 1208 AA; 125548 MW; B650B53AE5312B CRC64;
 Query Match 15.8%; Score 355; DB 16; Length 1208;
 Best Local Similarity 31.7%; Pred. No. 2,1e-12;
 Matches 126; Conservative 51; Mismatches 151; Indels 70; Gaps 17;
 QY 8 VKADVQNN-----FGLYGQGIIVAVADTGLDTRNDSSVHEAFRGKITALYALG 57
 DB 186 VEAADMAESNAQITRAAWDAGLTGDTGVTVAVLDTGVDT-----THPDLGRVSRKSF 239
 QY 58 RTNNANDPNHGTHVAGSVLG-----NATNKGMAPOANLVFQSIM-DSGGGLGLPANLQ 111
 DB 240 DEEEVADRNGHGHVTHVTSVGGSGAASDGTGERVAFGATLVAGKVLSDGAG-----SESQ 294
 QY 112 TLFQAYSAG--GARIHNSGAPVNGAYTTD--SRNVDDYVRKNDMTILFAAGNEGPGS 166
 DB 295 ILAGWEAARDVRARIVSVSLGS--TEASDGTDPMAEAVDTLSEETGALFVVAAGNTGAPS 353
 QY 167 GTISAPGTAKNAITVGATENLRFSGSYADNINHVAQFSSRGPT--DGRIKPDVMAPGTY 225
 DB 354 -SIGSGAADSALTVA-----VDSGDRAAFTSAGPRHGDNALKPDLAAPGVD 401
 QY 226 ILSARSSLAPDSSFVANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLK 285
 DB 402 IRAARSQAPGTGY-----YTSMSGTSMATPHVAGVALLAEQHPDWTGARLKDALMS 454
 QY 286 AA--LIAGAADVGLGPNQWQ--RVTLDKSLNVA-----VNETSPSTSOXA 332
 DB 455 TSEQLDASVYQAGRGVSPDVAAGVATATGSADLGFHRWPHDAPRPTKVTYTSNSD 514
 QY 333 TYSFT--AQAGKPLKSLVMSDA---PGSTASLTSLVND 366
 DB 515 TVELSLVRGAPAGVATLADTALTYPANGHTAATTGVD 552
 RESULT 15
 P95684
 ID P95684 PRELIMINARY; PRT; 1102 AA.
 AC P95684;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Subtilisin-like protease.
 OS Streptomyces albogriseolus.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S-3253;
 RX MEDLINE=97144528; PubMed=8990295;
 RA Suzuki M., Taguchi S., Yamada S., Kojima S., Miura K., Momose H.,
 RT "A novel member of the subtilisin-like protease family from
 RT Streptomyces albogriseolus.",
 RL J. Bacteriol. 179:430-438(1997).
 CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 DR EMBL; D83672; BAA12040.1; -.
 DR HSP; P00782; ZSBT.
 DR MEROPS; S08.069; -.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002860; GH_ENR.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF02012; ENR; 2.
 DR Pfam; PF00082; Peptidase S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydrolase; Protease; Serine protease.
 SQ SEQUENCE 1102 AA; 114128 MW; F9E4AD2590FE559E CRC64;
 Query Match 15.2%; Score 341.5; DB 2; Length 1102;
 Best Local Similarity 30.2%; Pred. No. 1,1e-11;
 Matches 132; Conservative 51; Mismatches 181; Indels 73; Gaps 16;
 QY 18 GLYGGQGIIVAVADTGLDTRNDSSVHEAFRGKITALYALGRNNANDPNHGTHVAGSVL 77
 DB 212 GYDGKGVKIAVLDTGVD-----ATHPDLKGQVTKNFSTAPTGDVYVGHGTHVASIAA 265
 QY 78 G-----NATNKGMAPOANLVFQSIMDSG--GGGLGLPANLQTLFQAYSAGARIHNSWG 130
 DB 266 GTGAQSKGIYGVAPGAKILNGKVLDDAGFGDDSGILAGMEWAAQ-----GADIVNMSLG 321
 QY 131 A-----PVNGAYTTDSNRVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAITVGA 183
 DB 322 GMDTPETDPLEAA-----VDKLSAEKGLFALFAAGNEGPOS--IGSPGSDSALTVA 372
 QY 184 TENLRFSGSYADNINHVAQFSSRGPT--TRDGRIPKPDVMAPGTYILSARSLAPDGSFWAN 242
 DB 373 -----VDDKDLADFSTGPRLDGAVKPDLTAPGVDITAAASAKGNDAKEVGE 421
 QY 243 HDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAGAADVGLGFPNG 302
 DB 422 KPAGYMTISGTSMATPHVAGAAALLKQHP-----WKYAEKGLALTASTKD--GKYTP-F 474
 QY 303 NOGWGVTLDKSLNVAFVNETSPLS-----TSQATYSFTQAAGKPLKISLV 349
 DB 475 EQSGRGVQDKATQTVIAEPVSLSGVQVQWPHADKPVTKKLTYNLRTEDVTTLKLTST 534
 QY 350 WSDAPGSTTASLTSLVNDLDTLTPANGTKYVGNDFAP--YDNNMDGRNNVENVFNAPO 407
 DB 535 ATGPKGAAPAGPFTLGASTLTVPANGTASV--DVTADTRLGAVDGTYSYVATGAGQ 592
 QY 408 S-----GTYIVVEQAYNV 420
 DB 593 SVETAAAREVERESYNV 609
 Search completed: March 31, 2004, 16:09:02
 Job time : 33.475 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 47.1304 Seconds
(without alignments)
2595.843 Million cell updates/sec

Title: US-09-985-689A-7
Perfect score: 2252
Sequence: 1 NDVARGIVKADVAQNNGFLY.....EVQAYNPVSPQTSFLAIHV 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|--------|---------------|--------|-------|--------------------|
| 1 | 2252 | 100.0 | 433 | 5 | AAM50086 Bacillus |
| 2 | 2252 | 100.0 | 641 | 2 | AAM99547 Bacillus |
| 3 | 2125.5 | 94.4 | 434 | 5 | AAM50080 Bacillus |
| 4 | 2125.5 | 94.4 | 640 | 2 | AAY17090 Bacillus |
| 5 | 2120.5 | 94.2 | 640 | 2 | AAY17091 Bacillus |
| 6 | 2118.5 | 94.1 | 434 | 5 | AAM50081 Bacillus |
| 7 | 2116.5 | 94.0 | 434 | 5 | AAM50085 Bacillus |
| 8 | 2110.5 | 93.7 | 639 | 2 | AAY17089 Bacillus |
| 9 | 2060.5 | 91.5 | 639 | 2 | AAY17087 An alkali |
| 10 | 2060.5 | 91.5 | 640 | 2 | AAY17088 An alkali |
| 11 | 2044 | 90.8 | 433 | 5 | AAM50084 Bacillus |
| 12 | 2040 | 90.6 | 433 | 5 | AAM50082 Bacillus |
| 13 | 2038 | 90.5 | 636 | 2 | AAM89548 Bacillus |
| 14 | 2033 | 90.3 | 433 | 5 | AAM50083 Bacillus |
| 15 | 2032 | 90.2 | 433 | 2 | AAR26274 Alkali-pr |
| 16 | 2032 | 90.2 | 433 | 2 | AAM61495 Modified |
| 17 | 2032 | 90.2 | 433 | 2 | AAM95698 Bacillus |
| 18 | 2032 | 90.2 | 433 | 3 | AAY69207 Amino aci |
| 19 | 2032 | 90.2 | 433 | 3 | AAY44619 Bacillus |
| 20 | 1949.5 | 86.6 | 434 | 5 | AAM50090 Bacillus |
| 21 | 1649 | 73.2 | 345 | 2 | AAM62230 Subtilase |
| 22 | 1649 | 73.2 | 345 | 2 | AAY21654 Subtilase |
| 23 | 432.5 | 19.2 | 659 | 2 | AAW24121 Thermococ |
| 24 | 432.5 | 19.2 | 659 | 2 | AAW94840 WO9856926 |
| 25 | 396 | 17.6 | 412 | 2 | AAW94836 Hyperther |

26 396 17.6 522 2 AAW241122 Pyrococcus
27 396 17.6 522 2 AAW94838 Hyperther
28 396 17.6 654 2 AAW24129 Pyrococcus
29 396 17.6 654 2 AAW94841 Hyperther
30 390 17.3 659 2 AAW24123 Protease.
31 369.5 16.4 545 4 ABB09483 T. yonsei
32 341.5 15.2 1079 4 AAB81180 Transglut
33 341.5 15.2 1079 6 ABU07391 Foreign P
34 327.5 14.5 520 2 AAW13666 Streptomy
35 327.5 14.5 734 2 AAW13667 Streptomy
36 327.5 14.5 823 2 AAW13668 DhpA-mel
37 323 14.3 1237 6 ABU11343 Protein e
38 309 13.7 806 2 AAR27481 RP-III re
39 306 13.6 903 2 AAR87007 Hyperther
40 306 13.6 1398 2 AAR87008 Protease.
41 306 13.6 1398 2 AAW24124 Pyrococcus
42 306 13.6 1398 2 AAW94839 WO9856926
43 297 13.2 519 6 ABP76735 Streptomy
44 297 13.2 19938 6 ABP76678 Streptomy
45 289 12.8 699 2 AAY08471 F. balust

ALIGNMENTS

RESULT 1
AAM50086
ID AAM50086 standard; protein: 433 AA.
XX
AC AAM50086;
XX
DT 12-AUG-2002 (first entry)
XX
DE Bacillus sp alkaline protease protein A-2 fragment.
XX
KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX
OS Bacillus sp.
XX
PN EP1209233-A2.
XX
PD 29-MAY-2002.
XX
PF 22-NOV-2001; 2001EP-00127851.
XX
PR 22-NOV-2000; 2000JP-00355166.
PR 12-APR-2001; 2001JP-00114048.
XX
PA (KAOS) KAO CORP.
XX
PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
PI Okuda M, Sasaki K;
XX
DR WPI; 2002-437518/47.
XX
PT New modified alkaline proteases useful in detergent compositions.
XX
PS Claim 5; Page 20-21; 25pp; English.
XX
CC This invention describes novel Bacillus sp. alkaline proteases useful in
CC detergent compositions, especially in laundry, bleaching or automatic
CC dishwasher detergents. The novel proteases have an increased detergency *
CC (34 ~ 38%) compared to prior art alkaline proteases (31 and 23%). This
CC sequence represents a fragment of the alkaline protease A-2 from Bacillus
CC sp NCIB12513 described in the method of the invention
XX
SQ Sequence 433 AA;

Query Match 100.0%; Score 2252; DB 5; Length 433;
Best Local Similarity 100.0%; Pred. No. 1.2e-160;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGFLYGGQIVAVADTGLTGRNDSNMHEAFRGKITALVLRGTN 60

Db 1 NDVARGIVKADVAQNFGYGGQIVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 60
Qy 61 NANDPNHGHTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 120
Db 61 NANDPNHGHTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 120
Qy 121 GARIHTNSWGAPVNGGTYTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
Db 121 GARIHTNSWGAPVNGGTYTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
Qy 181 VGATENLRPSFGSYADNINHVAFVNETSPSTSQKATYSFTAQAGKPLKISLVMSDAPGTTAS 240
Db 181 VGATENLRPSFGSYADNINHVAFVNETSPSTSQKATYSFTAQAGKPLKISLVMSDAPGTTAS 240
Qy 241 ANHDSKYAYMGGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAAIAGAADVGLGFP 300
Db 241 ANHDSKYAYMGGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAAIAGAADVGLGFP 300
Qy 301 NGNQGWGRVTLDKSLNVAFNWETSPSTSQKATYSFTAQAGKPLKISLVMSDAPGTTAS 360
Db 301 NGNQGWGRVTLDKSLNVAFNWETSPSTSQKATYSFTAQAGKPLKISLVMSDAPGTTAS 360
Qy 361 LTLVNDLDELVITAPNGTKVGNDFTPAYDNNWGRNVENVFINAPQSGTYTVEQAYNV 420
Db 361 LTLVNDLDELVITAPNGTKVGNDFTPAYDNNWGRNVENVFINAPQSGTYTVEQAYNV 420
Qy 421 PVSPQTFSLAIVH 433
Db 421 PVSPQTFSLAIVH 433

RESULT 2
AAW89547
ID AAW89547 standard; protein; 641 AA.
XX
AC AAW89547;
XX
DT 12-APR-1999 (first entry)
XX
DE Bacillus JPI70 protease.
XX
KW Protease; detergent; surfactant; leather processing; debittering;
KW flavour.
XX
OS Bacillus sp.
XX
FH Key Location/Qualifiers
FT Peptide 1..33
FT Region /note= "signal peptide"
FT Region 34..208
FT Protein /note= "prepro region"
FT Protein 209..641
FT Protein /note= "mature protein"
XX
PN WO9856927-A2.
XX
PN 17-DEC-1998.
XX
PF 09-JUN-1998; 98WO-US012005.
XX
PR 12-JUN-1997; 97US-00873479.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
XX
PI Sloma A, Christianson L;
XX
PI WPI; 1999-080908/07.
DR N-PSDB; AAV82382.
XX
PT Novel protease from Bacillus subtilis LC20 - useful in laundry and
PT dishwashing detergents and for leather processing.
XX

PS Claim 7; Page 53-54; 77pp; English.
XX
CC This is the amino acid sequence of a novel protease of Bacillus sp. JPI70
CC (NCIB 12513), as deduced from the nucleotide sequence of an isolated gene
CC (see AAV82382). The entire protein, including the signal peptide and
CC prepro region, has 77% identity to alkaline protease Y (see AAW89548)
CC from Bacillus. The invention provides vectors, recombinant host cells and
CC methods for the recombinant production of the protease. The protease is
CC used in laundry and dishwashing detergents, for institutional and
CC industrial cleaning, and for leather processing, as well as for
CC debittering and enhancing the degree of hydrolysis of protein
CC hydrolysates, for flavour development through hydrolysis of proteins.
CC degradation of undesired peptides and in enzymatic synthesis of peptides.
CC It has enhanced stability towards oxidation under alkaline conditions,
CC e.g. towards bleaching agents of the peroxy type. The invention also
CC provides mutant cells in which the protease activity is diminished. Such
CC cells can be used for the production of heterologous recombinant proteins
XX
SQ Sequence 641 AA;
Query Match 100.0%; Score 2252; DB 2; Length 641;
Best Local Similarity 100.0%; Pred. No. 2.1e-160;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 NDVARGIVKADVAQNFGYGGQIVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 60
Db 209 NDVARGIVKADVAQNFGYGGQIVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 268
Qy 61 NANDPNHGHTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 120
Db 269 NANDPNHGHTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 328
Qy 121 GARIHTNSWGAPVNGGTYTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
Db 329 GARIHTNSWGAPVNGGTYTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 388
Qy 181 VGATENLRPSFGSYADNINHVAFVNETSPSTSQKATYSFTAQAGKPLKISLVMSDAPGTTAS 240
Db 389 VGATENLRPSFGSYADNINHVAFVNETSPSTSQKATYSFTAQAGKPLKISLVMSDAPGTTAS 448
Qy 241 ANHDSKYAYMGGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAAIAGAADVGLGFP 300
Db 449 ANHDSKYAYMGGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAAIAGAADVGLGFP 508
Qy 301 NGNQGWGRVTLDKSLNVAFNWETSPSTSQKATYSFTAQAGKPLKISLVMSDAPGTTAS 360
Db 509 NGNQGWGRVTLDKSLNVAFNWETSPSTSQKATYSFTAQAGKPLKISLVMSDAPGTTAS 568
Qy 361 LTLVNDLDELVITAPNGTKVGNDFTPAYDNNWGRNVENVFINAPQSGTYTVEQAYNV 420
Db 569 LTLVNDLDELVITAPNGTKVGNDFTPAYDNNWGRNVENVFINAPQSGTYTVEQAYNV 628
Qy 421 PVSPQTFSLAIVH 433
Db 629 PVSPQTFSLAIVH 641
RESULT 3
AAW50080
ID AAW50080 standard; protein; 434 AA.
XX
AC AAW50080;
XX
DT 12-AUG-2002 (first entry)
XX
DE Bacillus sp KSM-KP43 alkaline protease protein fragment.
XX
KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX
OS Bacillus sp.
XX
PN EPI209233-A2.
XX

PD 29-MAY-2002.
 XX
 XX 22-NOV-2001; 2001EP-00127851.
 XX
 XX 22-NOV-2000; 2000JP-00355166.
 PR 12-APR-2001; 2001JP-00114048.
 XX
 XX (KAOS) KAO CORP.
 XX
 XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 PI Okuda M, Saeki K;
 XX
 XX WPI; 2002-437518/47.
 XX
 XX New modified alkaline proteases useful in detergent compositions.
 PT
 XX Claim 1; Page 10-11; 25pp; English.
 XX
 CC This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency &
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease KP43 from
 CC Bacillus sp strain KSM-KP43 which is used to create the modified protease
 CC represented in AAM50090
 XX
 XX Sequence 434 AA;
 SQ
 Query Match 94.4%; Score 2125.5; DB 5; Length 434;
 Best Local Similarity 93.5%; Pred. No. 3.9e-151;
 Matches 406; Conservative 19; Mismatches 8; Indels 1; Gaps 1;
 QY 1 NDVARGIVKADVAQNNGFLYGQGIIVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
 Db 1 NDVARGIVKADVAQSSIGLYGQGIIVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGN-ATNKGMAPOANLVFQSIIMDSGGGLGGLPANLQTLFSQAYS 119
 Db 61 NANDPNHGHTHVAGSVLGNSTNKGMAPOANLVFQSIIMDSGGGLGGLPSNIQTLSQAYS 120
 QY 120 AGARIHTNSWGAVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179
 Db 121 AGARIHTNSWGAANGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 180
 QY 180 TVGATENLRPSFGSYADNINHVAQFSRGPTRDGRIKPDVMAFGTYILSARSSLPDSSF 239
 Db 181 TVGATENLRPSFGSYADNINHVAQFSRGPTRDGRIKPDVMAFGTYILSARSSLPDSSF 240
 QY 240 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKRGVTTPKPSLLKAALIAGAADVGLGF 299
 Db 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKRGVTTPKPSLLKAALIAGAADVGLGF 300
 QY 300 PNGNQGWGRVTLDKSLNVAFNETSQKATYSFTAQAGKPLKISLVWSDAPGSTTA 359
 Db 301 PNGNQGWGRVTLDKSLNVAFNETSQKATYSFTAQAGKPLKISLVWSDAPGSTTA 360
 QY 360 SITLVNDLVLITAPNGTKYVGNFTAPYDNNWDCGRNVENVFINAPSGTYTVEVQAYN 419
 Db 361 SVTLVNDLVLITAPNGTKYVGNFTAPYDNNWDCGRNVENVFINAPSGTYTVEVQAYN 420
 QY 420 VFPVQPTFSLAIVH 433
 Db 421 VFPVQPTFSLAIVN 434
 RESULT 4
 AAY17090
 ID AAY17090 standard; protein; 640 AA.
 XX
 AC AAY17090;
 DT 20-MAR-2003 (revised)
 DT 21-JUL-1999 (first entry)

XX Bacillus alkaline protease.
 DE
 XX Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
 KW washing composition; oxidising agent.
 XX
 OS Bacillus sp.
 XX
 XX WO9918218-A1.
 XX
 XX 15-APR-1999.
 XX
 XX 07-OCT-1998; 98WO-JP004528.
 XX
 XX 07-OCT-1997; 97JP-00274570.
 XX
 XX (KAOS) KAO CORP.
 XX
 XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
 PI Shikata S, Nomura M;
 XX
 XX WPI; 1999-287736/27.
 DR N-PSDB; AAX37278.
 XX
 XX Alkali protease from Bacillus used in washing powders.
 PT
 XX
 XX Disclosure; Page 58-63; 71pp; Japanese.
 PS
 XX The invention relates to alkaline proteases produced by strains of
 CC Bacillus. The proteases ability to digest casein is not inhibited by
 CC oleic acid and they have a high stability to oxidising agents. The
 CC alkaline protease of the invention has the following properties: (a) it
 CC is active over the pH range 4-13 and has at least 80% of its optimum
 CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is
 CC stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d)
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has
 CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
 CC used as enzymes in washing compositions for use in automatic dishwashers
 CC and for washing clothes. The stability to oxidising agents allows the
 CC enzyme to be an effective component of washing compositions including
 CC bleaches. The present sequence represents an alkaline protease. (Updated
 CC on 20-MAR-2003 to correct DR field.)
 XX
 XX Sequence 640 AA;
 SQ
 Query Match 94.4%; Score 2125.5; DB 2; Length 640;
 Best Local Similarity 93.5%; Pred. No. 6.7e-151;
 Matches 406; Conservative 19; Mismatches 8; Indels 1; Gaps 1;
 QY 1 NDVARGIVKADVAQNNGFLYGQGIIVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
 Db 207 NDVARGIVKADVAQSSIGLYGQGIIVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 266
 QY 61 NANDPNHGHTHVAGSVLGN-ATNKGMAPOANLVFQSIIMDSGGGLGGLPANLQTLFSQAYS 119
 Db 267 NANDPNHGHTHVAGSVLGNSTNKGMAPOANLVFQSIIMDSGGGLGGLPSNLQTLFSQAYS 326
 QY 120 AGARIHTNSWGAVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179
 Db 327 AGARIHTNSWGAANGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 386
 QY 180 TVGATENLRPSFGSYADNINHVAQFSRGPTRDGRIKPDVMAFGTYILSARSSLPDSSF 239
 Db 387 TVGATENLRPSFGSYADNINHVAQFSRGPTRDGRIKPDVMAFGTYILSARSSLPDSSF 446
 QY 240 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKRGVTTPKPSLLKAALIAGAADVGLGF 299
 Db 447 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKRGVTTPKPSLLKAALIAGAADVGLGF 506
 QY 300 PNGNQGWGRVTLDKSLNVAFNETSQKATYSFTAQAGKPLKISLVWSDAPGSTTA 359
 Db 507 PNGNQGWGRVTLDKSLNVAFNETSQKATYSFTAQAGKPLKISLVWSDAPGSTTA 566

QY 360 SLTLVNDLVLVITAPNGTKYVGNDFTPAPYDNNWDCGNNVNFVINAPOSQTYTVEQAYN 419
 Db 567 SVTLVNDLVLVITAPNGTKYVGNDFTPAPYDNNWDCGNNVNFVINAPOSQTYTVEQAYN 626

QY 420 VPVSPQTFSLAIVH 433
 Db 627 VPVGPQTFSLAIVN 640

RESULT 5
 AAY17091
 ID AAY17091 standard; protein; 640 AA.
 XX
 AC AAY17091;
 XX
 DT 20-MAR-2003 (revised)
 DT 21-JUL-1999 (first entry)
 XX
 DE Bacillus alkaline protease.
 XX
 KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
 KW washing composition; oxidising agent.
 XX
 OS Bacillus sp.
 XX
 FN WO9918218-A1.
 XX
 PD 15-APR-1999.
 XX
 PF 07-OCT-1998; 98WO-JP004528.
 XX
 PR 07-OCT-1997; 97JP-00274570.
 XX
 PA (KAOS) KAO CORP.
 XX
 PI Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
 PI Shikata S, Nomura M;
 XX
 DR WPI; 1999-287736/27.
 DR N-PSDB; AAX37279.
 XX
 PT Alkali protease from Bacillus used in washing powders.
 XX
 PS Disclosure; Page 63-68; 71pp; Japanese.
 CC
 CC The invention relates to alkaline proteases produced by strains of
 CC Bacillus. The proteases ability to digest casein is not inhibited by
 CC oleic acid and they have a high stability to oxidising agents. The
 CC alkaline protease of the invention has the following properties: (a) it
 CC is active over the pH range 4-13 and has at least 80% of its optimum
 CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
 CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has
 CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
 CC used as enzymes in washing compositions for use in automatic dishwashers
 CC and for washing clothes. The stability to oxidising agents allows the
 CC enzyme to be an effective component of washing compositions including
 CC bleaches. The present sequence represents an alkaline protease. (Updated
 CC on 20-MAR-2003 to correct DR field.)
 XX
 SQ Sequence 640 AA;

Query Match 94.2%; Score 2120.5; DB 2; Length 640;
 Best Local Similarity 93.3%; Pred. No. 1.6e-150;
 Matches 405; Conservative 19; Mismatches 9; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNEGLYCGOIVAVADTGLDGTGNDSSMHEAFRGKITALVALGRIN 60
 Db 207 NDVARGIVKADVAQSSYGLYGGQGVAVADTGLDGTGNDSSMHEAFRGKITALVALGRIN 266

QY 61 NANDPNHGHTHVAGSVLGN-ATNKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFQAYS 119
 Db 267 NANDTNGHTHVAGSVLGNSTNGKMAPQANLVFQSIMDSGGGLGGLPSNLQTLFQAYS 326

QY 120 AGARIHTNSGAPVNGAYTTDSNVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179
 Db 327 AGARIHTNSGAAVNGAYTTDSNVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 386

QY 180 TVGATENLRPSFGSYADNINHVAQFSRGTDRGRIPKDVMAFGTYILSARSSLAPDSSF 239
 Db 387 TVGATENLRPSFGSYADNINHVAQFSRGTDRGRIPKDVMAFGTYILSARSSLAPDSSF 446

QY 240 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAAIAGAADVGLGF 299
 Db 447 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAAIAGAADVGLGF 506

QY 300 PNGNQGWGRVTLDKSLNVAFVNETSPLTSQKATYSFTAQAGKPLKISLVMSDAPGSTTA 359
 Db 507 PNGNQGWGRVTLDKSLNVAFVNETSPLTSQKATYSFTAQAGKPLKISLVMSDAPGSTTA 566

QY 360 SLTLVNDLVLVITAPNGTKYVGNDFTPAPYDNNWDCGNNVNFVINAPOSQTYTVEQAYN 419
 Db 567 SVTLVNDLVLVITAPNGTKYVGNDFTPAPYDNNWDCGNNVNFVINAPOSQTYTVEQAYN 626

QY 420 VPVSPQTFSLAIVH 433
 Db 627 VPVGPQTFSLAIVN 640

RESULT 6
 AAM50081
 ID AAM50081 standard; protein; 434 AA.
 XX
 AC AAM50081;
 XX
 DT 12-AUG-2002 (first entry)
 DE Bacillus sp KSM-KP9860 alkaline protease protein fragment.
 KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 OS Bacillus sp.
 XX
 PN EPI209233-A2.
 XX
 PD 29-MAY-2002.
 XX
 PF 22-NOV-2001; 2001EP-00127851.
 XX
 PR 22-NOV-2000; 2000JP-00355166.
 PR 12-APR-2001; 2001JP-00114048.
 XX
 PA (KAOS) KAO CORP.
 XX
 PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 PI Okuda M, Saeki K;
 XX
 DR WPI; 2002-437518/47.
 XX
 PT New modified alkaline proteases useful in detergent compositions.
 XX
 PS Claim 5; Page 12-13; 25pp; English.
 CC
 CC This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency %
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease KP9860 from
 CC Bacillus sp strain KSM-KP9860 described in the method of the invention

XX SQ Sequence 434 AA;

Query Match 94.1%; Score 2118.5; DB 5; Length 434;
 Best Local Similarity 93.3%; Pred. No. 1.3e-150;
 Matches 405; Conservative 18; Mismatches 10; Indels 1; Gaps 1;


```

XX  Disclosure; Page 53-58; 71pp; Japanese.
PS  The invention relates to alkaline proteases produced by strains of
XX  Bacillus. The proteases ability to digest casein is not inhibited by
CC  oleic acid and they have a high stability to oxidising agents. The
CC  alkaline protease of the invention has the following properties: (a) it
CC  is active over the pH range 4-13 and has at least 80% of its optimum
CC  activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC  stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC  its ability to digest casein is not inhibited by oleic acid; (e) it has
CC  molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC  used as enzymes in washing compositions for use in automatic dishwashers
CC  and for washing clothes. The stability to oxidising agents allows the
CC  enzyme to be an effective component of washing compositions including
CC  bleaches. The present sequence represents an alkaline protease. (Updated
CC  on 20-MAR-2003 to correct DR field.)
XX  SQ  Sequence 639 AA;

Query Match          93.7%; Score 2110.5; DB 2; Length 639;
Best Local Similarity 93.1%; Pred. No. 9e-150;
Matches 404; Conservative 18; Mismatches 11; Indels 1; Gaps 1;

QY  1 NDVARGIVKADVAQNFGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
DB  206 NDVARGIVKADVAQSSYGLYGGQIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 265

QY  61 NANDPNGHGHVAGSVLGN-ATNKGMAPOANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 119
DB  266 NANDTNGHGHVAGSVLGNATNKGMAPOANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 325

QY  120 AGARIHTNSGAPVNGAYTTDSRVNDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179
DB  326 AGARIHTNSGAPVNGAYTTDSRVNDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 385

QY  180 TVGATENLRPFSGYADNINHEVAFQSSRGPTDRGRIKPDVWAPGTIILGARS LAPDSGF 239
DB  386 TVGATENLRPFSGYADNINHEVAFQSSRGPTDRGRIKPDVWAPGTIILGARS LAPDSGF 445

QY  240 WANHDSKYAYNGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKALIAAGADVGLGF 299
DB  446 WANHDSKYAYNGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKALIAAGADVGLGF 505

QY  300 PNGNQGWGRVTLDKSLNVAFYNETSPILSTSQKATYSFTAQAGKPLKISLWSDAPGSTTA 359
DB  506 PNGNQGWGRVTLDKSLNVAFYNETSPILSTSQKATYSFTAQAGKPLKISLWSDAPGSTTA 565

QY  360 SLTLVNDLVLITAPNGTKYVGNDFTPAYDNNWDRNNVNFVFNAPQSGTYTVEVQAYN 419
DB  566 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWDRNNVNFVFNAPQSGTYTVEVQAYN 625

RESULT 9
AAV17087
ID  AAV17087 standard; protein; 639 AA.
XX  AC  AAV17087;
XX  DT  20-MAR-2003 (revised)
XX  DT  21-JUL-1999 (first entry)
XX  DE  An alkaline protease sequence from Bacillus species.
XX  KW  Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX  washing composition; oxidising agent.
XX  OS  Bacillus sp.
XX

```

```

PH  Key  Location/Qualifiers
FT  Misc-difference 1..639
FT  /note= "all residues indicated as Xaa are arbitrary amino
XX  acids"
XX  WO9918218-A1.
XX  15-APR-1999.
XX  07-OCT-1998; 98WO-JP004528.
XX  07-OCT-1997; 97JP-00274570.
XX  (KAOS ) KAO CORP.
XX  Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
XX  Shikata S, Nomura M;
XX  WPI; 1999-287736/27.
XX  N-PSDB; AAX37277.
XX  Alkali protease from Bacillus used in washing powders.
XX  Claim 3; Page 47-50; 71pp; Japanese.
XX  The invention relates to alkaline proteases produced by strains of
XX  Bacillus. The proteases ability to digest casein is not inhibited by
XX  oleic acid and they have a high stability to oxidising agents. The
XX  alkaline protease of the invention has the following properties: (a) it
XX  is active over the pH range 4-13 and has at least 80% of its optimum
XX  activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
XX  stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
XX  its ability to digest casein is not inhibited by oleic acid; (e) it has
XX  molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
XX  used as enzymes in washing compositions for use in automatic dishwashers
XX  and for washing clothes. The stability to oxidising agents allows the
XX  enzyme to be an effective component of washing compositions including
XX  bleaches. The present sequence represents an alkaline protease of the
XX  invention. (Updated on 20-MAR-2003 to correct DR field.)
XX  SQ  Sequence 639 AA;

Query Match          91.5%; Score 2060.5; DB 2; Length 639;
Best Local Similarity 91.7%; Pred. No. 5.1e-146;
Matches 398; Conservative 12; Mismatches 23; Indels 1; Gaps 1;

QY  1 NDVARGIVKADVAQNFGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
DB  206 NDVARGIVKADVAQSSYGLYGGQIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 265

QY  61 NANDPNGHGHVAGSVLGN-ATNKGMAPOANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 119
DB  266 NANDTNGHGHVAGSVLGNATNKGMAPOANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 325

QY  120 AGARIHTNSGAPVNGAYTTDSRVNDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179
DB  326 AGARIHTNSGAPVNGAYTTDSRVNDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 385

QY  180 TVGATENLRPFSGYADNINHEVAFQSSRGPTDRGRIKPDVWAPGTIILSARSLAPDSGF 239
DB  386 TVGATENLRPFSGYADNINHEVAFQSSRGPTDRGRIKPDVWAPGTIILSARSLAPDSGF 445

QY  240 WANHDSKYAYNGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKALIAAGADVGLGF 299
DB  446 WANHDSKYAYNGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKALIAAGADVGLGF 505

QY  300 PNGNQGWGRVTLDKSLNVAFYNETSPILSTSQKATYSFTAQAGKPLKISLWSDAPGSTTA 359
DB  506 PNGNQGWGRVTLDKSLNVAFYNETSPILSTSQKATYSFTAQAGKPLKISLWSDAPGSTTA 565

QY  360 SLTLVNDLVLITAPNGTKYVGNDFTPAYDNNWDRNNVNFVFNAPQSGTYTVEVQAYN 419
DB  566 SVTLVNDLVLITAPNGTKYVGNDFTPAYDNNWDRNNVNFVFNAPQSGTYTVEVQAYN 625

```

QY 420 VPVSPQTFSLAIVH 433
Db 626 VPVGPQXFLAIVN 639

RESULT 10
AAV17088
ID AAY17088 standard; protein; 640 AA.
XX AAY17088;
AC AAY17088;
XX 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX An alkaline protease sequence from Bacillus species.
DE
XX Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
KW washing composition; oxidising agent.
XX
OS Bacillus sp.
PH Key Location/Qualifiers
FT Misc-difference 1..640
FT /note= "all residues indicated as Xaa are arbitrary amino acids"
FT
XX
XX WO9918218-A1.
XX
XX 15-APR-1999.
XX
XX 07-OCT-1998; 98WO-JP004528.
XX
XX 07-OCT-1997; 97JP-00274570.
XX
XX (XAOs) KAO CORP.
XX
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;
XX
XX WPI; 1999-287736/27.
DR N-PSDB; AAX37278.
XX
XX Alkali protease from Bacillus used in washing powders.
PT
XX
PS Claim 3; Page 50-53; 71pp; Japanese.
XX
XX The invention relates to alkaline proteases produced by strains of Bacillus. The proteases ability to digest casein is not inhibited by oleic acid and they have a high stability to oxidising agents. The alkaline protease of the invention has the following properties: (a) it is active over the pH range 4-13 and has at least 80% of its optimum activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d) its ability to digest casein is not inhibited by oleic acid; (e) it has molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be used as enzymes in washing compositions for use in automatic dishwashers and for washing clothes. The stability to oxidising agents allows the enzyme to be an effective component of washing compositions including bleaches. The present sequence represents an alkaline protease of the invention. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 640 AA;
Query Match 91.5%; Score 2060.5; DB 2; Length 640;
Best Local Similarity 91.7%; Pred. No. 5.1e-146;
Matches 398; Conservative 12; Mismatches 23; Indels 1; Gaps 1;
QY 1 NDVARGIVKADVAQNFGYGGQIVAVADTGLDGTGRNDSMHEAPRGKITALVALGRTN 60
Db 207 NDVARGIVKADVAQSSYGLYGGQIVAVADTGLDGTGRNDSMHEAPRGKITALVALGRTN 266
QY 61 NANDPENGHTHVAGSVLGN-ATNKGMAPQANLVFQSIMDSXGGLGGLPSNLQTLFQAYS 119

Db 267 NANDTNGHTHVAGSVLGN-ATNKGMAPQANLVFQSIMDSXGGLGGLPSNLQTLFQOAXS 326
QY 120 AGARIHTNSWGAPVNGAYTTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179
Db 327 AGARIHTNSWGAAVNGAYTTTDSRNVDYVRKNDMTILFAAGNEXPNGGTISAPGTAKNAI 386
QY 180 TVGATENLRPSFGSYADNINHVAFQSRGPTRDGRIPKPDVWAPGTIYLSARSSLAPDSSF 239
Db 387 TVGATENLRPSFGSYADNINHVAFQSRGPTKDGRIKPDVWAPGTIYLSARSSLAPDSSF 446
QY 240 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKALITAGAADYGLGF 299
Db 447 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKALITAGAADXGLGY 506
QY 300 PNGNQGWGRVTLDKSLNVAFVNETSPLSTSQKATYSFTAQAGKPLKISLVWSDAPGSTTA 359
Db 507 PNGNQGWGRVTLDKSLNVAYVNESSXLSSTQKATYFTATAGKPLKISLVWSDAPASTTA 566
QY 360 SLTLVNDLVLITAPNGTKYVGNDFTPAYDNNNDGRNNVENFINAPQSGTITVEVOAYN 419
Db 567 SVTLVNDLVLITAPNGTKYVGNDFXPPXXNNDGRNNVENFINXPSQSGTITVEVOAYN 626
QY 420 VPVSPQTFSLAIVH 433
Db 627 VPVGPQXFLAIVN 640

RESULT 11
AAV50084
ID AAM50084 standard; protein; 433 AA.
XX
AC AAM50084;
XX
XX 12-AUG-2002 (first entry)
XX
XX Bacillus sp SD-521 (FERM BP-11162) alkaline protease protein fragment.
XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX
XX Bacillus sp.
XX EP1209233-A2.
XX
XX 29-MAY-2002.
XX
XX 22-NOV-2001; 2001EP-00127851.
XX
XX 22-NOV-2000; 2000JP-00355166.
PR 12-APR-2001; 2001JP-00114048.
XX
XX (KAOS) KAO CORP.
XX
XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
PI Okuda M, Saeki K;
XX
XX WPI; 2002-437518/47.
XX
XX New modified alkaline proteases useful in detergent compositions.
PT
XX Claim 5; Page 16-18; 25pp; English.
XX
XX This invention describes novel Bacillus sp. alkaline proteases useful in detergent compositions, especially in laundry, bleaching or automatic dishwasher detergents. The novel proteases have an increased detergency (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This sequence represents a fragment of the alkaline protease SD-521 from Bacillus sp strain SD-521 (FERM BP-11162) described in the method of the CC invention
XX
SQ Sequence 433 AA;
Query Match 90.8%; Score 2044; DB 5; Length 433;

| | | |
|--|--|-----|
| Best Local Similarity 89.1%, Pred. No. 5.2e-145; | | |
| Matches | Conservative 24; Mismatches 23; Indels 0; Gaps 0 | |
| QY | 1 NDVARGIVKADVAQNNGFLYGQGGIVAVADTGLDTGENDSSMHEAFRGKITIYALGRTN | 60 |
| Db | 1 NDVARGIVKADVAQNNGFLYGQGGIVAVADTGLDTGENDSSMHEAFRGKITIYALGRTN | 60 |
| QY | 61 NANDPNHGHTHVAGSVLGNATKGMAPAQNLVFOSIMDSGGIGGGHPANLQTLFSQAQYA | 120 |
| Db | 61 NANDPNHGHTHVAGSVLGNALNKGMAQAQNLVFOSIMDSGGIGGLGSLNUNTLFSQAQNA | 120 |
| QY | 121 GARIHTNSWGAPVNGAYTTDSRNVDYVYRKNDMTLFLAAGNEGSGGTISAPGTAKNAIT | 180 |
| Db | 121 GARIHTNSWGAPVNGAYTANSRQVDVEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT | 180 |
| QY | 181 VGATENLRPSPFGSYADNINHVAAQFSSRGQPTRDGRIKPDVMAPGTYILSARSSILAPDSSFW | 240 |
| Db | 181 VGATENYRPSFSGSLADNPNHIAQFSSRGATFDGRIKPDVTAPGFTLSARSSILAPDSSFW | 240 |
| QY | 241 ANHDSKIAYMGTSMTATIVAGNVAQLEBHFVKRGVTPKPSLLKAALTAGAADVGLGFP | 300 |
| Db | 241 ANYNSKIAYMGTSMTATIVAGNVAQKREHFPIKRGITTPKPSLIKKAALTAGATDVLGLVP | 300 |
| QY | 301 NGNQMGRYVTLDKSLNVAFVNETSPLSQKATYSFTAQAGKPKLISLYMSDAPGSGTTAS | 360 |
| Db | 301 SGDQMGRYVTLDKSLNVAFVNEATALATGQKATYSFQAQAGKPKLISLVWDAPGSGTTAS | 360 |
| QY | 361 LTLVNDLDLAVITAPNGTKVYGNDETAPYDNNWNGRNNVNFVNAPOSGTYTVEVOAYNV | 420 |
| Db | 361 YTLVNDLDLAVITAPNGQKIVGNDFSPYDNNWNGRNNVNFVNAPOSGTYTTEVOAYNV | 420 |
| QY | 421 PVSPQTFSLAIYH 433 | |
| Db | 421 PSGPQFSLAIYH 433 | |

RESULT 12
AAM50082
ID AAM50082 standard: protein: 433 AA.

```

CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
CC sequence represents a fragment of the alkaline protease E-1 from Bacillus
CC sp strain D6- (FERM-P1592) described in the method of the invention
XX
SQ Sequence 433 AA;
      Query Match      90.6%; Score 2040; DB 5; Length 433;
      Best Local Similarity 88.9%; Pred. No. le-14;
      Matches 385; Conservative 24; Mismatches 24; Indels 0; Gaps 0;
      1 NDVARGIVKADVAGNNPGLYGGQIVAVADTGLDTCGRNDSMHEAFRGKITALYALGRTN 60
      Db 1 NDVARGIVKADVAGNNYGLYGGQIVAVADTGLDTCGRNDSMHEAFRGKITALYALGRTN 60
      61 NANDPNHGHTHVAGSVLGNATNKGMAPOANLVFQSIMDSGGGLGGTIPANLQTLFSQAYSA 120
      Db 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSGGGLGGTIPSNLTILFSQAWN 120
      121 GARLHTNSWGAQVNGAVYTTDSRNVDVVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
      Db 121 GARLHTNSWGAQVNGAVTANSRQVDEIVRRKNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
      181 VGATENLRPSFGSYADNINHVAFSSRGPRTDGRIKPDVNPAGPTIYLISARSSLAPDSSF 240
      Db 181 VGATENYRPSFGSLADNPNIHAFSSRGATEDGRIKPDVTPAGPTFILSARSSLAPDSSF 240
      241 ANHDSKAYVGGTSMATPIVAGNVAQUREHFVKRGVTPRPSSLKAAIAGAADVGLGFP 300
      Db 241 ANVNSKAYVGGTSMATPIVAGNVAQUREHFVKRGVTPRPSSLKAAIAGATDVGLGYP 300
      301 NGNOGWRVTLDKSLNAVFNETSPLGTSOKATYSETAQAGKPLKISLVNSDAPGSTTAS 360
      Db 301 SGDQGWGRVTLDKSLNAVYNEATALTGGQATYVSFQTQAGKPLKISLVNWDAPGSTTAS 360
      361 LTVLNDLDLVITAPNGTKYVGNDFIAPYDNNWGRNNVENVFVINAPOSGTYTVEVQAYNV 420
      Db 361 YTLVNDLDLVITAPNGQKYVGNDFSYPYDNNWGRNNVENVFVINAPOSGTYTVEVQAYNV 420
      421 PVSPQTFSLAIVH 433
      Db 421 PSGPQRFSLAIVH 433

```

| | |
|-----------|---|
| RESULT 13 | |
| AAW89548 | |
| ID | AAW89548 standard; protein; 636 AA. |
| XX | |
| AC | AAW89548; |
| XX | |
| DT | 12-APR-1999 (first entry) |
| XX | |
| DE | Bacillus sp. alkaline protease Y. |
| XX | |
| KW | Alkaline protease Y; detergent; surfactant; leather processing; |
| KW | debittering; flavour. |
| XX | |
| OS | Bacillus sp. |
| XX | |
| FN | WO9856927-A2. |
| XX | |
| PD | 17-DEC-1998. |
| XX | |
| PF | 09-JUN-1998; 98WO-US012005. |
| XX | |
| PR | 12-JUN-1997; 97US-00873479. |
| XX | |
| PA | (NOVO) NOVO NORDISK BIOTECH INC. |
| XX | |
| PI | Sloma A, Christianson L; |
| XX | |
| XX | WPI; 1999-080908/07. |
| DR | |
| PT | Novel protease from Bacillus subtilis LC20 - useful in laundry |

PT dishwashing detergents and for leather processing.
 XX Claim 3; Page 55-56; 77pp; English.
 PS
 XX This is the amino acid sequence of a Bacillus sp. alkaline protease Y
 CC that is said to have good alkali and surfactant resistance and improved
 CC detergency. It shows 77% identity to a newly isolated protease (see
 CC AA089547) of Bacillus sp. JF170 (NCIB 12513). The invention provides
 CC vectors, recombinant host cells and methods for the recombinant
 CC production of such proteases. The protease are used in laundry and
 CC dishwashing detergents, for institutional and industrial cleaning, and
 CC for leather processing, as well as for debittering and enhancing the
 CC degree of hydrolysis of protein hydrolysates, for flavour development
 CC through hydrolysis of proteins, degradation of undesired peptides and in
 CC enzymatic synthesis of peptides. They have enhanced stability towards
 CC oxidation under alkaline conditions, e.g. towards bleaching agents of the
 CC peroxy type. The invention also provides mutant cells in which the
 CC protease activity is diminished. Such cells can be used for the
 CC production of heterologous recombinant proteins
 XX
 SQ Sequence 636 AA;
 Query Match 90.5%; Score 2038; DB 2; Length 636;
 Best Local Similarity 88.9%; Pred. No. 2.5e-144;
 Matches 385; Conservative 25; Mismatches 23; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVQANNFGLYGQGIIVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
 DB 204 NDVARGIVKADVQANNFGLYGQGIIVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 263
 QY 61 NANDPNHGHTHVAGSVLGNATNKGMAPQANLVFQSIMDSGGGLGLPANLQTLFSQAYSA 120
 DB 264 NASDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSGGGLGLPSNLTLFSQAWNA 323
 QY 121 GARIHTNSWAPVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
 DB 324 GARIHTNSWAPVNGAYTANSRQVDEYVRNDMTVLFAAGNEGPNSTISAPGTAKNAIT 383
 QY 181 VGATENLRPSFGSVADNINHVAFSSRGPTDRGRIKPDVWAPGTIILSARSSLAPDSSF 240
 DB 384 VGATENYRPSFGSIADPNPHIAQFSSRGATDRGRIKPDVWAPGTIILSARSSLAPDSSF 443
 QY 241 ANHDSKYAYMGGTSMATPIVAGNVAQLREHFVNKRGVTPKPSLLKAALIAGADVGLGFP 300
 DB 444 ANYSKYAYMGGTSMATPIVAGNVAQLREHFVNKRGVTPKPSLLKAALIAGADVGLGYP 503
 QY 301 NGNOGWRVTLDKSLNVAFNVNETSPSTOKATYSFTAQAGKPLKISLVMSDAPGSTTAS 360
 DB 504 SGDOGWRVTLDKSLNVAFNVNETSPSTOKATYSFTAQAGKPLKISLVMSDAPGSTTAS 563
 QY 361 LTLVNDLDLVTAPNGTKYVGNDFTPAYDNNMDGRNNVNFVFNAPQSGTYTVEVQAYNV 420
 DB 564 YTLVNDLDLVTAPNGTKYVGNDFSAFYDNNMDGRNNVNFVFNAPQSGTYTVEVQAYNV 623
 QY 421 PVSQPTFSLAIVH 433
 DB 624 PSQPQFSLAIVH 636
 RESULT 14
 AAM50083
 ID AAM50083 standard; protein; 433 AA.
 XX
 AC AAM50083;
 XX
 XX 12-AUG-2002 (first entry)
 DT
 DE Bacillus sp Y-(FERM BP-1029) alkaline protease protein fragment.
 XX
 KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 OS Bacillus sp.
 XX

PN EPI209233-A2.
 XX
 PD 29-MAY-2002.
 XX
 PF 22-NOV-2001; 2001EP-00127851.
 XX
 PR 22-NOV-2000; 2000JP-00355166.
 PR 12-APR-2001; 2001JP-00114048.
 XX
 PA (KAOS) KAO CORP.
 XX
 PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 PI Okuda M, Saeki K;
 XX
 DR WPI; 2002-437518/47.
 XX
 PT New modified alkaline proteases useful in detergent compositions.
 XX
 PS Claim 5; Page 15-16; 25pp; English.
 XX
 CC This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency %
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease Ya from Bacillus
 CC sp strain Y-(FERM BP-1029) described in the method of the invention
 XX
 SQ Sequence 433 AA;
 Query Match 90.3%; Score 2033; DB 5; Length 433;
 Best Local Similarity 88.7%; Pred. No. 3.5e-144;
 Matches 384; Conservative 25; Mismatches 24; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVQANNFGLYGQGIIVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVQANNFGLYGQGIIVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGNATNKGMAPQANLVFQSIMDSGGGLGLPANLQTLFSQAYSA 120
 DB 61 NASDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSGGGLGLPSNLTLFSQAWNA 120
 QY 121 GARIHTNSWAPVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
 DB 121 GARIHTNSWAPVNGAYTANSRQVDEYVRNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
 QY 181 VGATENLRPSFGSVADNINHVAFSSRGPTDRGRIKPDVWAPGTIILSARSSLAPDSSF 240
 DB 181 VGATENYRPSFGSIADPNPHIAQFSSRGATDRGRIKPDVWAPGTIILSARSSLAPDSSF 240
 QY 241 ANHDSKYAYMGGTSMATPIVAGNVAQLREHFVNKRGVTPKPSLLKAALIAGADVGLGFP 300
 DB 241 ANYSKYAYMGGTSMATPIVAGNVAQLREHFVNKRGVTPKPSLLKAALIAGADVGLGYP 300
 QY 301 NGNOGWRVTLDKSLNVAFNVNETSPSTOKATYSFTAQAGKPLKISLVMSDAPGSTTAS 360
 DB 301 NGDOGWRVTLDKSLNVAFNVNETSPSTOKATYSFTAQAGKPLKISLVMSDAPGSTTAS 360
 QY 361 LTLVNDLDLVTAPNGTKYVGNDFTPAYDNNMDGRNNVNFVFNAPQSGTYTVEVQAYNV 420
 DB 361 YTLVNDLDLVTAPNGTKYVGNDFSAFYDNNMDGRNNVNFVFNAPQSGTYTVEVQAYNV 420
 QY 421 PVSQPTFSLAIVH 433
 DB 421 PSQPQFSLAIVH 433
 RESULT 15
 AAR26274
 ID AAR26274 standard; protein; 433 AA.
 XX
 AC AAR26274;
 XX
 DT 05-FEB-1993 (first entry)

XX Alkali-protease Ya enzyme.
DE
XX
XX Alkali resistance; surface active agent resistance; detergency improver.
XX
XX Bacillus sp. Y.
OS
XX
XX JP04197182-A.
PN
XX
XX 16-JUL-1992.
FD
XX
XX 28-NOV-1990; 90JP-00327110.
XX
XX 28-NOV-1990; 90JP-00327110.
XX
XX (LLOY) LION CORP.
PA
XX
XX WPI; 1992-288440/35.
DR
XX N-PSDB; AAQ27516.
DR
XX
XX DNA coding alkali-protease Ya enzyme - has good alkali and surfactant
PT resistance and improves detergency.
PT
XX
XX Claim 2; Page 1; 17pp; Japanese.
PS
XX
XX The sequence is that of alkali-protease Ya enzyme which can be used in
CC the recombinant production of Ya enzyme. Ya enzyme is excellent in alkali
CC resistance and surface active agent resistance and improves detergency
CC
XX
XX Sequence 433 AA;
SQ

Query Match 90.2%; Score 2032; DB 2; Length 433;
Best Local Similarity 88.7%; Pred. No. 4.1e-144;
Matches 384; Conservative 25; Mismatches 24; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNFGLYGGOIIVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
DB 1 NDVARGIVKADVAQNNFGLYGGOIIVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60

QY 61 NADPNHGHTHAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFQAYSA 120
DB 61 NADPNHGHTHAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFQAYSA 120

QY 121 GARIHTNSGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNALT 180
DB 121 GARIHTNSGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNALT 180

QY 181 VGATENLRPSFGSYADNINHVAQFSSRGPTDGRKIPDYMPCYILSARSLAPDSFFW 240
DB 181 VGATENLRPSFGSYADNINHVAQFSSRGPTDGRKIPDYMPCYILSARSLAPDSFFW 240

QY 241 ANHDSKYAYNGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALTAGADYGLGFP 300
DB 241 ANYNSKYAYNGGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALTAGADYGLGYP 300

QY 301 NGQGWGRVTLDKSLNVAFYNETSPLSTSQATYSFTAQAGPLKISLVMSDAPGSTTAS 360
DB 301 SGQGWGRVTLDKSLNVAFYNETSPLSTSQATYSFTAQAGPLKISLVMSDAPGSTTAS 360

QY 361 LTVNDLDDLVITAPNGTKYVGNDFTPADYDNNWGRNNVNFVFNAPQSGTYTVEQAYNV 420
DB 361 YTLVNDLDDLVITAPNGTKYVGNDFTPADYDNNWGRNNVNFVFNAPQSGTYTVEQAYNV 420

QY 421 PVSQPTFSLAIVH 433
DB 421 PSQPTFSLAIVH 433

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:59:39 ; Search time 13.9892 Seconds
(without alignments)
1597.947 Million cell updates/sec

Title: US-09-985-689A-7
Perfect score: 2252
Sequence: 1 NDVARGIVKADVAQNFGLY.....EVQAYNPVPSQTFSLAIVH 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A-COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B-COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A-COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B-COMB.pep:*
5: /cgn2_6/prodata/2/iaa/6C-COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------------|
| 1 | 2252 | 100.0 | 641 | 2 | US-08-873-479-42 |
| 2 | 2125.5 | 94.4 | 640 | 4 | US-09-509-814A-6 |
| 3 | 2120.5 | 94.2 | 640 | 4 | US-09-509-814A-8 |
| 4 | 2110.5 | 93.7 | 639 | 4 | US-09-509-814A-4 |
| 5 | 2060.5 | 91.5 | 639 | 4 | US-09-509-814A-1 |
| 6 | 2060.5 | 91.5 | 640 | 4 | US-09-509-814A-2 |
| 7 | 2032 | 90.2 | 433 | 4 | US-09-104-623A-4 |
| 8 | 2032 | 90.2 | 433 | 4 | US-09-019-532-4 |
| 9 | 2032 | 90.2 | 433 | 4 | US-08-338-746-4 |
| 10 | 2032 | 90.2 | 635 | 2 | US-08-873-479-43 |
| 11 | 1649 | 73.2 | 345 | 4 | US-09-512-251A-10 |
| 12 | 1649 | 73.2 | 345 | 4 | US-09-515-150A-10 |
| 13 | 1649 | 73.2 | 345 | 4 | US-09-196-281-13 |
| 14 | 432.5 | 19.2 | 659 | 3 | US-08-894-818B-1 |
| 15 | 432.5 | 19.2 | 659 | 4 | US-09-445-472-12 |
| 16 | 396 | 17.6 | 412 | 4 | US-09-445-472-1 |
| 17 | 396 | 17.6 | 522 | 3 | US-08-894-818B-3 |
| 18 | 396 | 17.6 | 522 | 4 | US-09-445-472-4 |
| 19 | 396 | 17.6 | 654 | 3 | US-08-894-818B-35 |
| 20 | 396 | 17.6 | 654 | 4 | US-09-445-472-16 |
| 21 | 390 | 17.3 | 659 | 3 | US-08-894-818B-5 |
| 22 | 327.5 | 14.5 | 520 | 3 | US-09-000-016-7 |
| 23 | 327.5 | 14.5 | 520 | 4 | US-09-514-340-7 |
| 24 | 327.5 | 14.5 | 734 | 3 | US-09-000-016-4 |
| 25 | 327.5 | 14.5 | 734 | 4 | US-09-514-340-4 |
| 26 | 327.5 | 14.5 | 823 | 3 | US-09-000-016-2 |
| 27 | 327.5 | 14.5 | 823 | 4 | US-09-514-340-2 |

| | | | | | | |
|----|-------|------|------|---|-------------------|-------------------|
| 28 | 306 | 13.6 | 903 | 1 | US-08-750-532-1 | Sequence 1, Appl |
| 29 | 306 | 13.6 | 1398 | 1 | US-08-750-532-9 | Sequence 9, Appl |
| 30 | 306 | 13.6 | 1398 | 3 | US-08-894-818B-8 | Sequence 8, Appl |
| 31 | 306 | 13.6 | 1398 | 4 | US-09-445-472-6 | Sequence 6, Appl |
| 32 | 275 | 12.3 | 237 | 1 | US-08-750-532-18 | Sequence 18, Appl |
| 33 | 271.5 | 12.1 | 418 | 1 | US-09-866-921A-2 | Sequence 2, Appl |
| 34 | 264.5 | 11.7 | 418 | 2 | US-08-873-479-44 | Sequence 44, Appl |
| 35 | 257.5 | 11.4 | 269 | 1 | US-07-706-691G-4 | Sequence 4, Appl |
| 36 | 257.5 | 11.4 | 269 | 2 | US-08-254-021-4 | Sequence 4, Appl |
| 37 | 257.5 | 11.4 | 269 | 3 | US-08-618-446-4 | Sequence 4, Appl |
| 38 | 257.5 | 11.4 | 269 | 3 | US-08-980-135-4 | Sequence 4, Appl |
| 39 | 257.5 | 11.4 | 269 | 4 | US-09-585-798-4 | Sequence 4, Appl |
| 40 | 256.5 | 11.4 | 269 | 1 | US-08-566-369-11 | Sequence 11, Appl |
| 41 | 256.5 | 11.4 | 269 | 3 | US-09-074-331-11 | Sequence 11, Appl |
| 42 | 256.5 | 11.4 | 269 | 5 | PCT-US95-01937-11 | Sequence 11, Appl |
| 43 | 255.5 | 11.3 | 269 | 1 | US-07-706-691G-5 | Sequence 5, Appl |
| 44 | 255.5 | 11.3 | 269 | 1 | US-08-254-021-5 | Sequence 5, Appl |
| 45 | 255.5 | 11.3 | 269 | 2 | US-08-618-446-5 | Sequence 5, Appl |

ALIGNMENTS

RESULT 1
US-08-873-479-42
; Sequence 42, Application US/08873473
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Lyne, Christian
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58917010 No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-873-479-42

Query Match 100.0%; Score 2252; DB 2; Length 641;
Best Local Similarity 100.0%; Pred. No. 1.8e-174;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNFGLYGQQQIVAVDTGTDGNDSSMHAFKKTALVALQRTN 60
|||||

Db 209 NDVARGIVKADVAQNNGFLYGGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRN 268
 Qy 61 NANDPNHGHTHVAGSVLGNATNKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 120
 Db 269 NANDPNHGHTHVAGSVLGNATNKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 328
 Qy 121 GARHTNSWGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 180
 Db 329 GARHTNSWGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 388
 Qy 181 VGATENLRPSFGSYADNINHVAFVNETSPSTSOXATYSFTAQAGKPLKISLVMSDAPGSTTA 240
 Db 389 VGATENLRPSFGSYADNINHVAFVNETSPSTSOXATYSFTAQAGKPLKISLVMSDAPGSTTA 448
 Qy 241 ANHDSKYAYMGSGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALAGAADVGLGPF 300
 Db 449 ANHDSKYAYMGSGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALAGAADVGLGPF 508
 Qy 301 NGNQGWGRVTLDKSLNVAFAVNETSPSTSOXATYSFTAQAGKPLKISLVMSDAPGSTTA 360
 Db 509 NGNQGWGRVTLDKSLNVAFAVNETSPSTSOXATYSFTAQAGKPLKISLVMSDAPGSTTA 568
 Qy 361 LTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGRNNVENVFINAPQSGTYTVEVOAYN 420
 Db 569 LTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGRNNVENVFINAPQSGTYTVEVOAYN 628
 Qy 421 PVSPQTFSLAIVH 433
 Db 629 PVSPQTFSLAIVH 641

RESULT 2

US-09-509-814A-6

; Sequence 6, Application US/09509814A

; Patent No. 6376227

; GENERAL INFORMATION:

; APPLICANT: TAKAIWA, MIKIO

; APPLICANT: OKUDA, MITSUYOSHI

; APPLICANT: SAKI, KATSUHIISA

; APPLICANT: KUBOTA, HIROMI

; APPLICANT: HITOMI, JUN

; APPLICANT: KAGEYAMA, YASUSHI

; APPLICANT: SHIKATA, SHITSUM

; APPLICANT: NOMURA, MASAFUMI

; TITLE OF INVENTION: ALKALINE PROTEASE

; FILE REFERENCE: 0327-0832-0PCT

; CURRENT APPLICATION NUMBER: US/09/509,814A

; PRIOR FILING DATE: 2000-04-06

; PRIOR APPLICATION NUMBER: PCT/JP98/04528

; PRIOR FILING DATE: 1998-10-07

; PRIOR FILING DATE: 1997-06-08

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 6

; LENGTH: 640

; TYPE: PRT

; ORGANISM: Bacillus sp.

US-09-509-814A-6

Query Match

Best Local Similarity 94.4%; Score 2125.5; DB 4; Length 640;

Matches 406; Conservative 19; Mismatches 8; Indels 1; Gaps 1;

Qy 1 NDVARGIVKADVAQNNGFLYGGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRN 60
 Db 207 NDVARGIVKADVAQSSYGLYGGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRN 266
 Qy 61 NANDPNHGHTHVAGSVLGN-ATNKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 119
 Db 267 NANDPNHGHTHVAGSVLGN-ATNKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 326
 Qy 120 AGARIHTNSWGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179

Db 327 AGARIHTNSWGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 386
 Qy 180 TVGATENLRPSFGSYADNINHVAFVNETSPSTSOXATYSFTAQAGKPLKISLVMSDAPGSTTA 239
 Db 387 TVGATENLRPSFGSYADNINHVAFVNETSPSTSOXATYSFTAQAGKPLKISLVMSDAPGSTTA 446
 Qy 240 WANHDSKYAYMGSGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALAGAADVGLGPF 299
 Db 447 WANHDSKYAYMGSGTSMATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALAGAADVGLGPF 506
 Qy 300 PNGNQGWGRVTLDKSLNVAFAVNETSPSTSOXATYSFTAQAGKPLKISLVMSDAPGSTTA 359
 Db 507 PNGNQGWGRVTLDKSLNVAFAVNETSPSTSOXATYSFTAQAGKPLKISLVMSDAPGSTTA 566
 Qy 360 SLTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGRNNVENVFINAPQSGTYTVEVOAYN 419
 Db 567 SLTLVNDLVLITAPNGTKYVGNDFTPAYDNNWGRNNVENVFINAPQSGTYTVEVOAYN 626
 Qy 420 PVSPQTFSLAIVH 433
 Db 627 PVSPQTFSLAIVH 640

RESULT 3

US-09-509-814A-8

; Sequence 8, Application US/09509814A

; Patent No. 6376227

; GENERAL INFORMATION:

; APPLICANT: TAKAIWA, MIKIO

; APPLICANT: OKUDA, MITSUYOSHI

; APPLICANT: SAKI, KATSUHIISA

; APPLICANT: KUBOTA, HIROMI

; APPLICANT: HITOMI, JUN

; APPLICANT: KAGEYAMA, YASUSHI

; APPLICANT: SHIKATA, SHITSUM

; APPLICANT: NOMURA, MASAFUMI

; TITLE OF INVENTION: ALKALINE PROTEASE

; FILE REFERENCE: 0327-0832-0PCT

; CURRENT APPLICATION NUMBER: US/09/509,814A

; PRIOR FILING DATE: 2000-04-06

; PRIOR APPLICATION NUMBER: PCT/JP98/04528

; PRIOR FILING DATE: 1998-10-07

; PRIOR APPLICATION NUMBER: JP 9-274570

; PRIOR FILING DATE: 1997-06-08

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 8

; LENGTH: 640

; TYPE: PRT

; ORGANISM: Bacillus sp.

US-09-509-814A-8

Query Match

Best Local Similarity 94.2%; Score 2120.5; DB 4; Length 640;

Matches 405; Conservative 19; Mismatches 9; Indels 1; Gaps 1;

Qy 1 NDVARGIVKADVAQNNGFLYGGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRN 60
 Db 207 NDVARGIVKADVAQSSYGLYGGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRN 266
 Qy 61 NANDPNHGHTHVAGSVLGN-ATNKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 119
 Db 267 NANDPNHGHTHVAGSVLGN-ATNKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 326
 Qy 120 AGARIHTNSWGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 179
 Db 327 AGARIHTNSWGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAI 386
 Qy 180 TVGATENLRPSFGSYADNINHVAFVNETSPSTSOXATYSFTAQAGKPLKISLVMSDAPGSTTA 239
 Db 387 TVGATENLRPSFGSYADNINHVAFVNETSPSTSOXATYSFTAQAGKPLKISLVMSDAPGSTTA 446

```
QY 240 WANHSKYAYMGTSMTATVAGNVAOLBEHFVKNRGVTPKPSLLKRALIAGAADVGLGF 299
Db 447 WANHSKYAYMGTSMTATVAGNVAOLBEHFVKNRGVTPKPSLLKRALIAGAADVGLGY 506
QY 300 PNGNQGWGRVTLDKSLNVAFAVNETSPSTSQKATYSFTAQAGKPLKISLWMSDAPGSTTA 359
Db 507 PNGNQGWGRVTLDKSLNVAFAVNETSPSTSQKATYSFTAQAGKPLKISLWMSDAPGSTTA 566
QY 360 SLTLVNDLIDLITAPNGTKYVGNDFTPAYDNWGDGRNNVNFNAPQSGTYYTVEVOAYN 419
Db 567 SVTLVNDLIDLITAPNGTKYVGNDFTPAYDNWGDGRNNVNFNAPQSGTYYTVEVOAYN 625
QY 420 VPVSPQTFSLAIYH 433
Db 627 VPVGPQNFSLAIYH 640

RESULT 4
US-09-509-814A-4
; Sequence 4, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-OPCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.
;
Query Match 93.7%; Score 2110.5; DB 4; Length 639;
Best Local Similarity 93.1%; Pred. No. 5.4e-163;
Matches 404; Conservative 18; Mismatches 11; Indels 1; Gaps 1;
QY 1 NDVARGIVKADVAQNFGLYGQGGIVAVADTGLDGRNDSMHEAPRGKITALYALGRTN 60
Db 206 NDVARGIVKADVAQSYGLYGQGGIVAVADTGLDGRNDSMHEAPRGKITALYALGRTN 265
QY 61 NANDPNGHGTHTVAGSVLGN-ATNKGMAPOANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 119
Db 266 NANDPNGHGTHTVAGSVLGNATNKGMAPOANLVFQSIMDSGGGLGGLPANLQTLFSQAYS 325
QY 120 AGARIHNSWGAQVNGAYTTDSRNVDDYVRKNDMTILPAAGNPGSGTISAPGTAKAI 179
Db 326 AGARIHNSWGAQVNGAYTTDSRNVDDYVRKNDMTILPAAGNPGSGTISAPGTAKAI 385
QY 180 TVGATENLRPFSGYADNINHAQFSSRGPTDRGRKPKDVMAGTIVLSARSLSAPDSFSF 239
Db 386 TVGATENLRPFSGYADNINHAQFSSRGPTDRGRKPKDVMAGTIVLSARSLSAPDSFSF 445
QY 240 WANHSKYAYMGTSMTATVAGNVAOLBEHFVKNRGVTPKPSLLKRALIAGAADVGLGF 299
Db 446 WANHSKYAYMGTSMTATVAGNVAOLBEHFVKNRGVTPKPSLLKRALIAGAADVGLGY 505
QY 300 PNGNQGWGRVTLDKSLNVAFAVNETSPSTSQKATYSFTAQAGKPLKISLWMSDAPGSTTA 359
Db 506 PNGNQGWGRVTLDKSLNVAFAVNETSPSTSQKATYSFTAQAGKPLKISLWMSDAPGSTTA 565
```

```
QY 360 SLTLVNDLIDLITAPNGTKYVGNDFTPAYDNWGDGRNNVNFNAPQSGTYYTVEVOAYN 419
Db 566 SVTLVNDLIDLITAPNGTKYVGNDFTPAYDNWGDGRNNVNFNAPQSGTYYTVEVOAYN 625
QY 420 VPVSPQTFSLAIYH 433
Db 626 VPVGPQNFSLAIYH 639

RESULT 5
US-09-509-814A-1
; Sequence 1, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUMI
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-OPCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.
;
NAME/KEY: misc feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (128)..(128)
```

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid

NAME/KEY: misc feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1

Query Match 91.5%; Score 2060.5; DB 4; Length 639;
Best Local Similarity 91.7%; Pred. No. 6.2e-159;
Matches 398; Conservative 12; Mismatches 23; Indels 1; Gaps 1;
Qy 1 NDVARGIVKADVAQNNGFLYGGQIVAVADTGLDTRNDSSMHEAPRGKITALYALGRTN 60
Db 206 NDVARGIVKADVAQSSVLYGGQIVAVADTGLDTRNDSSMHEAPRGKITALYALGRTN 265
Qy 61 NANDPNCHGTHVAGSVLGN-ATNKGMAPQANLVFQSIMDSGGGLGGLPANLOTLPQOAYS 119
Db 266 NANDTNGHGHVAGSVLGNXGTXNKGMAPQANLVFQSIMDSXGGLGGLPSNLQTLFQOAXS 325
Qy 120 AGARIHTNSWGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEPGSGGTISAPGTAKNAI 179
Db 326 AGARIHTNSWGAAVNGAYTTDSRNVDYVRKNDMTILFAAGNEPXNGGTISAPGTAKNAI 385
Qy 180 TVGATENLRPSFGSYADNINHVAFSGRGTGRKIPDVMAPGTVILSARSSLAPDSSF 239
Db 386 TVGATENLRPSFGSYADNINHVAFSGRGTGRKIPDVMAPGTVILSARSSLAPDSSF 445
Qy 240 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTTPKPSLLKAALIAGAADVGLGF 299
Db 446 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKNRGVTTPKPSLLKAALIAGAADXGLGY 505
Qy 300 PNGQHGGRVTLDKSLNVAFNETSPLSTQKATYFTAGQAGKPLKISLWSDAPGSTTA 359
Db 506 PNGQHGGRVTLDKSLNVAFNETSPLSTQKATYFTAGQAGKPLKISLWSDAPGSTTA 565
Qy 360 SLTLVNDLDLIVITAPNGTKYVGNDFTPAYDNNWDRNNVNFINAPQSGTYYTVEVOAYN 419
Db 566 SVTLVNDLDLIVITAPNGTKYVGNDFTPAYDNNWDRNNVNFINAPQSGTYYTVEVOAYN 625
Qy 420 VPVSPQTFSLAIVH 433
Db 626 VPVGPQXFLAIVN 639

RESULT 6
US-09-509-814A-2
Sequence 2, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUW
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A

;; CURRENT FILING DATE: 2000-04-06
;; PRIOR APPLICATION NUMBER: PCT/JP98/04528
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: JP 9-274570
;; PRIOR FILING DATE: 1997-06-08
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 2
;; LENGTH: 640
;; TYPE: PRT
;; ORGANISM: Bacillus sp.
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (3)..(3)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (24)..(24)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (30)..(30)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (33)..(33)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (47)..(47)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (48)..(48)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (54)..(54)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (71)..(71)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (75)..(75)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (90)..(90)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (103)..(103)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (106)..(106)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (129)..(129)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (131)..(131)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (132)..(132)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (133)..(133)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (134)..(134)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (147)..(147)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (149)..(149)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (161)..(161)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature

;; LOCATION: (166)..(166)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (173)..(173)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (184)..(184)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (188)..(188)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (189)..(189)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (190)..(190)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (195)..(195)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (287)..(287)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (307)..(307)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (325)..(325)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (370)..(370)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (432)..(432)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (502)..(502)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (532)..(532)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (542)..(542)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (585)..(585)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (592)..(592)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (593)..(593)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (595)..(595)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (596)..(596)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (597)..(597)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (612)..(612)
;; OTHER INFORMATION: Xaa is any amino acid
;; NAME/KEY: misc_feature
;; LOCATION: (633)..(633)
;; OTHER INFORMATION: Xaa is any amino acid
;; US-09-509-814A-2

Query Match 91.5%; Score 2060.5; DB 4; Length 640;
Best Local Similarity 91.7%; Pred. No. 6.2e-159;
Matches 398; Conservative 12; Mismatches 23; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNGFLYGGQIIVAVADTGLDTRNDSSMHEAFRGKITIYALGRTN 60
DB 1 NDVARGIVKADVAQNNGFLYGGQIIVAVADTGLDTRNDSSMHEAFRGKITIYALGRTN 60
QY 61 NADPENGHTHVAGSVLGNATKGNAPQANLVFQSIIMDSGGGLGGLPANLQTLFSQAYSA 120
DB 61 NASDPENGHTHVAGSVLGNATKGNAPQANLVFQSIIMDSGGGLGGLPANLQTLFSQAYSA 120
QY 121 GARIHTNSWGPVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
DB 121 GARIHTNSWGPVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
QY 181 VGATENLRPSFSGYADNHNHVAQFSRGPTRDGRIPKDVMAFGTITLSARSSSLAPDSSFW 240
DB 181 VGATENLRPSFSGYADNHNHVAQFSRGPTRDGRIPKDVMAFGTITLSARSSSLAPDSSFW 240
QY 241 ANHDSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAGAADVGLGYP 300
DB 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAGAADVGLGYP 300
QY 301 NGNQGRVTLDKSLNVAFVNETSPSTSQKATYFTQAQAGKPLKISLVMSDAPGSTTAS 360
DB 301 SGOQGRVTLDKSLNVAFVNETSPSTSQKATYFTQAQAGKPLKISLVMSDAPGSTTAS 360
QY 361 LTLVNDLVLITAPNGTKVGNDFTPYDNNWGRNVNFINAPQSGTYTVEVOAYNV 420
DB 361 YTLVNDLVLITAPNGTKVGNDFTPYDNNWGRNVNFINAPQSGTYTVEVOAYNV 420
QY 421 PVSQTFSLAIVH 433
DB 421 PSQPQFSLAIVH 433

RESULT 9
US-09-338-746-4
; Sequence 4, Application US/09338746
; Patent No. 6638526
; GENERAL INFORMATION:
; APPLICANT: Deussen, Heinz-Josef
; APPLICANT: Olsen, Arne A.
; APPLICANT: Fatum, Tine M.
; APPLICANT: Roggen, Erwin L.
; TITLE OF INVENTION: A Polypeptide-Polymer Conjugate
; FILE REFERENCE: 5619.200-US
; CURRENT APPLICATION NUMBER: US/09/338,746
; CURRENT FILING DATE: 1999-06-23
; EARLIER FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: 60/091,461
; EARLIER FILING DATE: 1998-07-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-338-746-4

Query Match 90.2%; Score 2032; DB 4; Length 433;
Best Local Similarity 88.7%; Pred. No. 7.3e-157;
Matches 384; Conservative 25; Mismatches 24; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQNNGFLYGGQIIVAVADTGLDTRNDSSMHEAFRGKITIYALGRTN 60
DB 1 NDVARGIVKADVAQNNGFLYGGQIIVAVADTGLDTRNDSSMHEAFRGKITIYALGRTN 60
QY 61 NADPENGHTHVAGSVLGNATKGNAPQANLVFQSIIMDSGGGLGGLPANLQTLFSQAYSA 120
DB 61 NASDPENGHTHVAGSVLGNATKGNAPQANLVFQSIIMDSGGGLGGLPANLQTLFSQAYSA 120
QY 121 GARIHTNSWGPVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180

DB 121 GARIHTNSWGPVNGAYTTANSRQVDEYVRNNDMTVLFAGNEGPNSTISAPGTAKNAIT 180
QY 181 VGATENLRPSFSGYADNHNHVAQFSRGPTRDGRIPKDVMAFGTITLSARSSSLAPDSSFW 240
DB 181 VGATENLRPSFSGYADNHNHVAQFSRGPTRDGRIPKDVMAFGTITLSARSSSLAPDSSFW 240
QY 241 ANHDSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAGAADVGLGYP 300
DB 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALIAGAADVGLGYP 300
QY 301 NGNQGRVTLDKSLNVAFVNETSPSTSQKATYFTQAQAGKPLKISLVMSDAPGSTTAS 360
DB 301 SGOQGRVTLDKSLNVAFVNETSPSTSQKATYFTQAQAGKPLKISLVMSDAPGSTTAS 360
QY 361 LTLVNDLVLITAPNGTKVGNDFTPYDNNWGRNVNFINAPQSGTYTVEVOAYNV 420
DB 361 YTLVNDLVLITAPNGTKVGNDFTPYDNNWGRNVNFINAPQSGTYTVEVOAYNV 420
QY 421 PVSQTFSLAIVH 433
DB 421 PSQPQFSLAIVH 433

RESULT 10
US-08-873-479-43
; Sequence 43, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Lynne, Christanson
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 58917010 No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agria, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-873-479-43

Query Match 90.2%; Score 2032; DB 2; Length 635;
Best Local Similarity 88.7%; Pred. No. 1.3e-156;
Matches 384; Conservative 25; Mismatches 24; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQNNGFLYGGQIIVAVADTGLDTRNDSSMHEAFRGKITIYALGRTN 60
DB 203 NDVARGIVKADVAQNNGFLYGGQIIVAVADTGLDTRNDSSMHEAFRGKITIYALGRTN 262

QY 61 NANDPNHGTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 120
Db 263 NASDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAWNA 322
QY 121 GARIHTNSGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
Db 323 GARIHTNSGAPVNGAYTANSQVDEYVRNNDMTILFAAGNEGPGSGTISAPGTAKNAIT 382
QY 181 VGATENLRPSFGSYADNINHVAQFSRSGPTRDRGRIKPDVMAPGTIILSARSSLAPDSFW 240
Db 383 VGATENLRPSFGSIADNPNHIAQFSRSGPTRDRGRIKPDVAPGTIILSARSSLAPDSFW 442
QY 241 ANHDSKYAYMGGTSMATPIVAGNVQALREHFVNKGVTPKPSLLKAAIAGAADVGLGFP 300
Db 443 ANYNSKYATMGGTSMATPIVAGNVQALREHFVNKGVTPKPSLLKAAIAGAADVGLGFP 502
QY 301 NGNQGWGRVTLDKSLNVAFVNETSPSTSQATYSFTAQAGKPLKISLVSDAPGSTTAS 360
Db 503 SGPQGWGRVTLDKSLNVAVYNEATATATQKATYSFQAQKPLKISLVTDAPGSTTAS 562
QY 361 LTLVNDLDLIVITAPNGTKYVGNDFTPADYNNWDGRNNVENFINAPQSGTYTVEQAYNV 420
Db 563 YTLVNDLDLIVITAPNGTKYVGNDFSYPDYNNWDGRNNVENFINAPQSGTYTVEQAYNV 622
QY 421 PVSQPTFSLAIVH 433
Db 623 PVSQPTFSLAIVH 635

RESULT 11
US-09-512-251A-10
; Sequence 10, Application US/09512251A
; Patent No. 6555355
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; APPLICANT: Andersen, Kim
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5349.204-US
; CURRENT APPLICATION NUMBER: US/09/512.251A
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Bacillus
US-09-512-251A-10

Query Match 73.2%; Score 1649; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 6.2e-126;
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQNPNFLYGGQGIIVAVADTGLDTRNDSSMHEAPRGKITALYALGRTN 60
Db 29 NDVARGIVKADVAQNPNFLYGGQGIIVAVADTGLDTRNDSSMHEAPRGKITALYALGRTN 88
QY 61 NANDPNHGTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 120
Db 89 NANDPNHGTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 148
QY 121 GARIHTNSGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
Db 149 GARIHTNSGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 208
QY 181 VGATENLRPSFGSYADNINHVAQFSRSGPTRDRGRIKPDVMAPGTIILSARSSLAPDSFW 240
Db 383 VGATENLRPSFGSIADNPNHIAQFSRSGPTRDRGRIKPDVAPGTIILSARSSLAPDSFW 442
QY 241 ANHDSKYAYMGGTSMATPIVAGNVQALREHFVNKGVTPKPSLLKAAIAGAADVGLGFP 300
Db 443 ANYNSKYATMGGTSMATPIVAGNVQALREHFVNKGVTPKPSLLKAAIAGAADVGLGFP 502
QY 301 NGNQGWGRVTLDKSLNVAFVNETSPSTSQATYSFTAQAGKPLKISLVSDAPGSTTAS 360
Db 503 SGPQGWGRVTLDKSLNVAVYNEATATATQKATYSFQAQKPLKISLVTDAPGSTTAS 562
QY 361 LTLVNDLDLIVITAPNGTKYVGNDFTPADYNNWDGRNNVENFINAPQSGTYTVEQAYNV 420
Db 563 YTLVNDLDLIVITAPNGTKYVGNDFSYPDYNNWDGRNNVENFINAPQSGTYTVEQAYNV 622
QY 421 PVSQPTFSLAIVH 433
Db 623 PVSQPTFSLAIVH 635

QY 301 NGNQGWGRVTLDKSLNV 317
Db 329 NGNQGWGRVTLDKSLNV 345
RESULT 12
US-09-515-150A-10
; Sequence 10, Application US/09515150A
; Patent No. 6558938
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; APPLICANT: Andersen, Kim
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5348.204-US
; CURRENT APPLICATION NUMBER: US/09/515.150A
; CURRENT FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Bacillus
US-09-515-150A-10

Query Match 73.2%; Score 1649; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 6.2e-126;
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQNPNFLYGGQGIIVAVADTGLDTRNDSSMHEAPRGKITALYALGRTN 60
Db 29 NDVARGIVKADVAQNPNFLYGGQGIIVAVADTGLDTRNDSSMHEAPRGKITALYALGRTN 88
QY 61 NANDPNHGTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 120
Db 89 NANDPNHGTHVAGSVLGNATKGMAPQANLVFQSIMDSGGGLGGLPANLQTLFSQAYSA 148
QY 121 GARIHTNSGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
Db 149 GARIHTNSGAPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 208
QY 181 VGATENLRPSFGSYADNINHVAQFSRSGPTRDRGRIKPDVMAPGTIILSARSSLAPDSFW 240
Db 209 VGATENLRPSFGSYADNINHVAQFSRSGPTRDRGRIKPDVAPGTIILSARSSLAPDSFW 268
QY 241 ANHDSKYAYMGGTSMATPIVAGNVQALREHFVNKGVTPKPSLLKAAIAGAADVGLGFP 300
Db 269 ANHDSKYAYMGGTSMATPIVAGNVQALREHFVNKGVTPKPSLLKAAIAGAADVGLGFP 328
QY 301 NGNQGWGRVTLDKSLNV 317
Db 329 NGNQGWGRVTLDKSLNV 345

RESULT 13
US-09-196-281-13
; Sequence 13, Application US/09196281A
; Patent No. 6605458
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter K.
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5435.200-US
; CURRENT APPLICATION NUMBER: US/09/196.281A
; CURRENT FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: 1332/97
; EARLIER FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13

LENGTH: 345
TYPE: PRT
ORGANISM: Bacillus
US-09-196-281-13

Query Match 73.2%; Score 1649; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 6.2e-126;
Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNFGLVGGQIVAVADTGLDGTGRNDSMHEAFRGKITALYALGRTN 60
DB 29 NDVARGIVKADVAQNNFGLVGGQIVAVADTGLDGTGRNDSMHEAFRGKITALYALGRTN 88
QY 61 NADPNHGHTHVAGSVLGNATKMGAPQANLVFQSIWDSGGGLGGLPANLQTLFQAYSA 120
DB 89 NADPNHGHTHVAGSVLGNATKMGAPQANLVFQSIWDSGGGLGGLPANLQTLFQAYSA 148
QY 121 GARHTNSWGPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
DB 149 GARHTNSWGPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 208
QY 191 VGATENLRPSGVSADNINHAQPSRSGPTRDGRIRKPDVMAPTIYLSARSSLAPDSSFW 240
DB 209 VGATENLRPSGVSADNINHAQPSRSGPTRDGRIRKPDVMAPTIYLSARSSLAPDSSFW 268
QY 241 ANHDSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALTAGADVGLGPP 300
DB 269 ANHDSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLLKAALTAGADVGLGPP 328
QY 301 NGNQGWGRVTLDKSLNV 317
DB 329 NGNQGWGRVTLDKSLNV 345

RESULT 14

US-08-894-818B-1
Sequence 1, Application US/08894818B
Patent No. 6261822

GENERAL INFORMATION:

APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MITTA, Masanori
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
APPLICANT: KATO, Ikunoshin

TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES

NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:

ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300

CITY: Washington

STATE: D.C.

COUNTRY: United States of America

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/08/894, 818B

Filing DATE: 20-MAY-1998

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03253

Filing DATE: 07-NOV-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 323285/1995

Filing DATE: 12-DEC-1995

ATTORNEY/AGENT INFORMATION:

NAME: Browdy, Roger L.

REGISTRATION NUMBER: 25,618

REFERENCE/DOCKET NUMBER: TAKAKURA=1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 628-5197

TELEFAX: (202) 737-3528

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 659 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-894-818B-1

Query Match 19.2%; Score 432.5; DB 3; Length 659;

Best Local Similarity 30.7%; Pred. No. 7.2e-27;

Matches 134; Conservative 62; Mismatches 156; Indels 85; Gaps 16;

QY 8 VKADVAQNNFGLVGGQIVAVADTGLDGTGRNDSMHEAFRGKITALY-ALGRTNNDPN 66
DB 145 IGADTVNLSLGYDGGGVVVAIVDTGIDAN-----HPDLKGVIGVYDAVNGRSTFYDDQ 198
QY 67 GHGTHVAGSVLG-----NATKMGAPQANLVFQSIM--DSGGGLGGLPANLQTLFQAYSA 120
DB 199 GHGTHVAGSVLG-----NATKMGAPQANLVFQSIM--DSGGGLGGLPANLQTLFQAYSA 258
QY 121 GARHTNSWGPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
DB 259 GIRVINLSLGSQSSDGTSLDQAVNNAWDAGIVVCAAGNSGPNVTYTVGSPAAASKVIT 318
QY 181 VGATENLRPSGVSADNINHAQPSRSGPTRDGRIRKPDVMAPTIYLSARSSLAPDSSFW 240
DB 319 VGA-----VDSNDNIASFSSRSGPTDGRLEKPEVVAQGVVDIIAPRAS---GTSMG 364
QY 241 ANHDSKYAYMGTSMTATPIVAG-NVAQLREHFVKNRGVTPK-PSLLKAALTAGA----- 292
DB 365 TPINDYTKASGTSMTATPHVSGVGAIIQAH-----PSWTPDKVKTALLETADIVAP 416
QY 293 ---ADVGLGFPNGNQGWGRVTLDKSL---NVAFNETSLSTSQKATYSFTQAQAKPLKI 346
DB 417 KEIADIAYGA-----GRVNVYKAIKYDDYAKLTFTGSAVDKGSATHTF----- 459
QY 347 SLVMSDAPGSGTASLTIV-----NDLDLVITAPNGTKYVGNDFATPYDNNWGRNNVNV 401
DB 460 -----DVSQATFVTATLYWDGTSSDIDLYLDENGNE-VDYSYTAAY-----GFKEV 505
QY 402 FINAQSGTYTVEQAY 418
DB 506 GYNTAGTWTWKVWSY 522

RESULT 15

US-09-445-472-12

Sequence 12, Application US/09445472

Patent No. 6358726

GENERAL INFORMATION:

APPLICANT: TAKAKURA, Hikaru

APPLICANT: MORISHITA, Mio

APPLICANT: SHIMOJO, Tomoko

APPLICANT: ASADA, Kiyozo

APPLICANT: KATO, Ikunoshin

TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE

FILE REFERENCE: TAKAKURA=6

CURRENT APPLICATION NUMBER: US/09/445, 472

PRIOR FILING DATE: 1999-12-06

PRIOR APPLICATION NUMBER: 151969/1997

PRIOR FILING DATE: 1997-06-10

NUMBER OF SEQ ID NOS: 33

SOFTWARE: Patent in version 3.0

SEQ ID NO 12

LENGTH: 659

TYPE: PRT

ORGANISM: Thermococcus celer

US-09-445-472-12

```
Query Match      19.2%; Score 432.5; DB 4; Length 659;
Best Local Similarity 30.7%; Pred. No. 7.2e-27;
Matches 134; Conservative 62; Mismatches 156; Indels 85; Gaps 16;

QY      8 VKADVAQNFGLYGGQIVAVADTGLDGRNDSMHEAFRGKITALLY-ALGRTNNANDPN 66
Db      145 IGADIVNNSLGVGGGVVAIVDTGIDAN-----HPDLKGVIGWYDAVNGRSTPYDDQ 198

QY      67 GHGTHVAGSVLG-----NATNKGMAFCANLVFOSIM--DSGGGLGGLPANLQTLFSQAYSA 120
Db      199 GHGTHVAGSVLG-----NATNKGMAFCANLVFOSIM--DSGGGLGGLPANLQTLFSQAYSA 120

QY      121 GARIHTNSWGPVNGAYTTDSRNVDYVRKNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
Db      259 GIRVINLSLGSQSSDGTDSLSCAVNNAWDAGIVVCVAGNSGPNYTVGSPAAASKVIT 318

QY      181 VGATENLRPSFGSYADNINHVAQFSGRPTDRGRIKPDVMAQPTIILSARSLAPDSSFW 240
Db      319 VGA-----VDSNDNIASFSGRPTDRGRIKPDVMAQPTIILSARSLAPDSSFW 240

QY      241 ANHDSKYAYMGGTSMATPIVAG-NVAQLREHFVKNRGVTPK--PSLLKAALIAGA----- 292
Db      365 TPINDYTKASGTSMATPHVSGVGAIILOAH-----PSWTPDKVKTALJETADIYAP 416

QY      293 ---ADYGLGFPNGNQGWRVTLDKSL---NVAFVNETSPLSTSQKATYSFTHAQKPLKI 346
Db      417 KEIADIAIGA-----GRVNVYKAICYDDYAKLTFTGSVADKGSATHTF----- 459

QY      347 SLVWSDAPGSTTASLTIV-----NDLDLVITAPNGTKYVGNDFTPYDNNWDRNNVENV 401
Db      460 -----DVSQATFTATLYWDTGSSDIDLILYDPNGNE-VDYSYTAYY-----GFEKV 505

QY      402 FINAPQSGTYTVEVQAY 418
Db      506 GYNNPTAGTWTVKWSY 522
```

Search completed: March 31, 2004, 16:11:57
Job time : 14.9892 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 47.1304 Seconds
(without alignments)
2595.843 Million cell updates/sec

Title: US-09-985-689A-4
Perfect score: 2260
Sequence: 1 NDVARGIVKADVAQNNGLY.....EVQAYNVSPGQRFSLAIVH 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_28Jan04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 2260 | 100.0 | 433 | 5 | AAM50083 Bacillus |
| 2 | 2247 | 99.4 | 433 | 2 | AAR26274 Alkali-pr |
| 3 | 2247 | 99.4 | 433 | 2 | AAR61495 Modified |
| 4 | 2247 | 99.4 | 433 | 2 | AAR95698 Bacillus |
| 5 | 2247 | 99.4 | 433 | 3 | AAY69207 Amino aci |
| 6 | 2247 | 99.4 | 433 | 3 | AAY44619 Bacillus |
| 7 | 2240 | 99.1 | 433 | 5 | AAM50084 Bacillus |
| 8 | 2238 | 99.0 | 433 | 5 | AAM89548 Bacillus |
| 9 | 2234 | 98.8 | 433 | 5 | AAM50082 Bacillus |
| 10 | 2033 | 90.0 | 433 | 5 | AAM50086 Bacillus |
| 11 | 2033 | 90.0 | 641 | 2 | AAR89547 Bacillus |
| 12 | 2005.5 | 88.7 | 434 | 5 | AAM50085 Bacillus |
| 13 | 1999.5 | 88.5 | 434 | 5 | AAM50081 Bacillus |
| 14 | 1991.5 | 88.1 | 639 | 2 | AAY17089 Bacillus |
| 15 | 1988.5 | 88.0 | 640 | 2 | AAY17091 Bacillus |
| 16 | 1987.5 | 87.9 | 434 | 5 | AAM50080 Bacillus |
| 17 | 1987.5 | 87.9 | 640 | 2 | AAY17090 Bacillus |
| 18 | 1941.5 | 85.9 | 639 | 2 | AAY17087 An alkali |
| 19 | 1941.5 | 85.9 | 640 | 2 | AAY17088 An alkali |
| 20 | 1830.5 | 81.0 | 434 | 5 | AAM50090 Bacillus |
| 21 | 1509 | 66.8 | 345 | 2 | AAR62230 Subtilase |
| 22 | 1509 | 66.8 | 345 | 2 | AAY21654 Subtilase |
| 23 | 438 | 19.4 | 659 | 2 | AAM24121 Thermococ |
| 24 | 438 | 19.4 | 659 | 2 | AAM94840 WO9856926 |
| 25 | 416.5 | 18.4 | 412 | 2 | AAM94836 Hyperther |

| | | | | | | |
|----|-------|------|-------|---|----------|------------|
| 26 | 416.5 | 18.4 | 522 | 2 | AAW24122 | Pyrococcus |
| 27 | 416.5 | 18.4 | 522 | 2 | AAW94838 | Hyperther |
| 28 | 416.5 | 18.4 | 654 | 2 | AAW24129 | Pyrococcus |
| 29 | 416.5 | 18.4 | 654 | 2 | AAW94841 | Hyperther |
| 30 | 395.5 | 17.5 | 659 | 2 | AAW24123 | Protease |
| 31 | 378.5 | 16.7 | 545 | 4 | ABSO9483 | T. yonsei |
| 32 | 339.5 | 15.0 | 1079 | 6 | ABSO1180 | Transglut |
| 33 | 339.5 | 15.0 | 1079 | 6 | ABU07391 | Foreign p |
| 34 | 337 | 14.9 | 520 | 2 | AAW13666 | Fragment |
| 35 | 337 | 14.9 | 734 | 2 | AAW13667 | Streptomy |
| 36 | 337 | 14.9 | 823 | 2 | AAW13668 | DhpA-mel |
| 37 | 334.5 | 14.8 | 1237 | 6 | ABU11343 | Protein e |
| 38 | 313 | 13.8 | 903 | 2 | AAR87007 | Hyperther |
| 39 | 313 | 13.8 | 1398 | 2 | AAR87008 | Protease |
| 40 | 313 | 13.8 | 1398 | 2 | AAW24124 | Pyrococcus |
| 41 | 313 | 13.8 | 1398 | 2 | AAW94839 | WO9856926 |
| 42 | 309 | 13.7 | 806 | 2 | AAR27481 | RP-III re |
| 43 | 300.5 | 13.3 | 580 | 7 | ADD24927 | Xanthomon |
| 44 | 298 | 13.2 | 519 | 6 | ABP76735 | Streptomy |
| 45 | 298 | 13.2 | 19338 | 6 | ABP76678 | Streptomy |

ALIGNMENTS

RESULT 1

AAM50083
ID AAM50083 standard; protein; 433 AA.

XX AAM50083;

DT 12-AUG-2002 (first entry)

XX DE Bacillus sp Y-(FERM BP-1029) alkaline protease protein fragment.

XX KW Alkaline protease; detergent; laundry; bleaching; dishwasher.

XX OS Bacillus sp.

XX PN EP12092333-A2.

XX PD 29-MAY-2002.

XX PF 22-NOV-2001; 2001EP-00127851.

XX PR 22-NOV-2000; 2000JP-00355166.

XX PR 12-APR-2001; 2001JP-00114048.

XX FA (KAOS) KAO CORP.

XX PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;

XX PI Okuda M, Saeki K;

XX DR WPI; 2002-437518/47.

XX PT New modified alkaline proteases useful in detergent compositions.

XX PS Claim 5; Page 15-16; 25pp; English.

XX CC This invention describes novel Bacillus sp. alkaline proteases useful in detergent compositions, especially in laundry, bleaching or automatic dishwasher detergents. The novel proteases have an increased detergency (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This sequence represents a fragment of the alkaline protease Ya from Bacillus sp strain Y-(FERM BP-1029) described in the method of the invention

XX SQ Sequence 433 AA;

Query Match 100.0%; Score 2260; DB 5; Length 433;

Best Local Similarity 100.0%; Pred. No. 5.5e-169;

Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGLYGQGVAVDTGLDTGSDSSMHEAFKGTALYALGRTN 60

```

Db 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGRNDSSMHEAPRGKITALVALGETN 60
QY 61 NASDPNGHGTTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
Db 61 NASDPNGHGTTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
QY 121 GARIHTNSGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTGISAPGTAKNAIT 180
Db 121 GARIHTNSGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTGISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRKPDVTAPGTFTLSARSLAPDSSFW 240
Db 181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRKPDVTAPGTFTLSARSLAPDSSFW 240
QY 241 ANVNSKYAVMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALTAGATDVGLGYP 300
Db 241 ANVNSKYAVMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALTAGATDVGLGYP 300
QY 301 NGDQGWGRVTLNKS LNVAVYNEATATLQKATYSFQAQAGKPLKISLVWTDAPGSTTAS 360
Db 301 NGDQGWGRVTLNKS LNVAVYNEATATLQKATYSFQAQAGKPLKISLVWTDAPGSTTAS 360
QY 361 YTLVNDLDLVTAPNGQKVGNDFSYPDNNWDRNNVENVFINAPQSGTYIIIEVQAYNV 420
Db 361 YTLVNDLDLVTAPNGQKVGNDFSYPDNNWDRNNVENVFINAPQSGTYIIIEVQAYNV 420
QY 421 PSQPQRFSLAIVH 433
Db 421 PSQPQRFSLAIVH 433

```

RESULT 2

```

AAR26274
ID AAR26274 standard; protein; 433 AA.
XX AC AAR26274;
XX DT 05-FEB-1993 (first entry)
XX DE Alkali-protease Ya enzyme.
XX KW Alkali resistance; surface active agent resistance; detergent improver.
XX OS Bacillus sp. Y.
XX PN JP04197182-A.
XX PD 16-JUL-1992.
XX PF 28-NOV-1990; 90JP-00327110.
XX PR 28-NOV-1990; 90JP-00327110.
XX PA (LLOY ) LION CORP.
XX DR WPI; 1992-288440/35.
XX DR N-PSDB; AAQ27516.
XX PT DNA coding alkali-protease Ya enzyme - has good alkali and surfactant
XX resistance and improves detergency.
XX PS Claim 2; Page 1; 17pp; Japanese.
XX CC The sequence is that of alkali-protease Ya enzyme which can be used in
XX the recombinant production of Ya enzyme. Ya enzyme is excellent in alkali
XX resistance and surface active agent resistance and improves detergency
XX SQ Sequence 433 AA;

```

```

Query March 99.4%; Score 2247; DB 2; Length 433;
Best Local Similarity 99.3%; Pred. NO. 5.7e-168;
Matches 430; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGRNDSSMHEAPRGKITALVALGETN 60
Db 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGRNDSSMHEAPRGKITALVALGETN 60
QY 61 NASDPNGHGTTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
Db 61 NASDPNGHGTTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
QY 121 GARIHTNSGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTGISAPGTAKNAIT 180
Db 121 GARIHTNSGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTGISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRKPDVTAPGTFTLSARSLAPDSSFW 240
Db 181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRKPDVTAPGTFTLSARSLAPDSSFW 240
QY 241 ANVNSKYAVMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALTAGATDVGLGYP 300
Db 241 ANVNSKYAVMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALTAGATDVGLGYP 300
QY 301 NGDQGWGRVTLNKS LNVAVYNEATATLQKATYSFQAQAGKPLKISLVWTDAPGSTTAS 360
Db 301 NGDQGWGRVTLNKS LNVAVYNEATATLQKATYSFQAQAGKPLKISLVWTDAPGSTTAS 360
QY 361 YTLVNDLDLVTAPNGQKVGNDFSYPDNNWDRNNVENVFINAPQSGTYIIIEVQAYNV 420
Db 361 YTLVNDLDLVTAPNGQKVGNDFSYPDNNWDRNNVENVFINAPQSGTYIIIEVQAYNV 420
QY 421 PSQPQRFSLAIVH 433
Db 421 PSQPQRFSLAIVH 433

```

RESULT 3

```

AAR61495
ID AAR61495 standard; protein; 433 AA.
XX AC AAR61495;
XX DT 06-NOV-1998 (first entry)
XX DE Modified Bacillus lion Y protease.
XX KW Bacillus lion Y protease; polyethylene glycol; PEG; soap;
XX KW methoxypolyethyleneglycol; mPEG; skin; hair care product; cosmetic;
XX OS lipstick; hair gel; sun oil; shampoo; hair dye; insect repellent.
XX PS Bacillus sp.
XX PF Key Location/Qualifiers
XX FT Modified-site 1..433
XX FT /note= "The enzyme is modified by methoxypolyethyleneglycol
XX FT molecules covalently attached to the N-terminal amino
XX FT group and to fourteen unspecified amino groups of lysine
XX FT residues present on the surface of the enzyme"
XX PN WO9830682-A1.
XX PD 16-JUL-1998.
XX PF 12-JAN-1998; 98WO-DK0000015.
XX PR 10-JAN-1997; 97DK-000000038.
XX PR 25-JUN-1997; 97DK-000000754.
XX PA (NOVO ) NOVO-NORDISK AS.
XX PI Olsen AA, Prento A;
XX WPI; 1998-399132/34.
XX PT New enzyme modified by attachment of many polymeric molecules - useful in

```

PT skin and hair care products, has reduced tendency to cause sensitisation
 PT and increased stability.

XX Claim 16; Page 44-45; 56pp; English.

XX The present sequence represents the *Bacillus* lion Y protease. The
 CC invention claims for enzymes covalently modified on their surface by the
 CC attachment of a large number of small polymeric molecules, e.g.
 CC polyethylene glycols (PEG). The polymeric molecules are coupled to the N-
 CC terminal amino group and the amino groups of lysine residues found on the
 CC surface of the enzyme. In the example given, the lion Y protease was
 CC modified using methoxypolyethyleneglycol (mPEG) as the polymeric molecule.
 CC The N-terminal amino group and the amino groups of the fourteen lysine
 CC residues present on the surface of the lion Y protease were modified.
 CC Modification of the enzymes increases the stability and/or reduces the
 CC sensitising potential (allergenicity) of the enzyme, without
 CC significantly reducing enzymatic activity. Also, using a large number of
 CC relatively small polymeric molecules, rather than a few very large ones,
 CC provides a more even effect with reduced activity loss. The modified
 CC enzymes are claimed to be useful as components of a wide range of skin
 CC and hair care products, e.g. soaps, cosmetics, creams, gels, lipsticks,
 CC hair gels, sun oils, shampoos, hair dyes, insect repellants, etc

XX Sequence 433 AA;

Query Match 99.4%; Score 2247; DB 2; Length 433;

Best Local Similarity 99.3%; Pred. No. 5.7e-168;
 Matches 430; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIKADVAAQNNYGLYGGQGVAVADTGLTGRNDSMHEAFRGKITALYALGRTN 60

DB 1 NDVARGIKADVAAQNNYGLYGGQGVAVADTGLTGRNDSMHEAFRGKITALYALGRTN 60

QY 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGPSNLNLFQAWNA 120

DB 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGPSNLNLFQAWNA 120

QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVNNDMTTLFAAGNEGPNSTGTSAPGTAKNAIT 180

DB 121 GARIHTNSWGAPVNGAYTANSRQVDEYVNNDMTTLFAAGNEGPNSTGTSAPGTAKNAIT 180

QY 181 VGATENYRPSFGSIADNPNHIAQSSRGATRDGRIPKDVTPAGTFTLSARSLAPDSSF 240

DB 181 VGATENYRPSFGSIADNPNHIAQSSRGATRDGRIPKDVTPAGTFTLSARSLAPDSSF 240

QY 241 ANYNSKYAVMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300

DB 241 ANYNSKYAVMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300

QY 301 NGDQGWGRVTLNKSINVAVNEATATGCKATYSFQAQKPLKISLVWTDAPGSTTAS 360

DB 301 NGDQGWGRVTLNKSINVAVNEATATGCKATYSFQAQKPLKISLVWTDAPGSTTAS 360

QY 361 YTLVNDLDELVITAPNGQKYGNDFSYPYDNNWGRNNVENVFINAPQSGTYIIIEVQAYNV 420

DB 361 YTLVNDLDELVITAPNGQKYGNDFSYPYDNNWGRNNVENVFINAPQSGTYIIIEVQAYNV 420

QY 421 PSGPQRFSLAIVH 433

DB 421 PSGPQRFSLAIVH 433

RESULT 4

AAW95698

ID AAW95698 standard; protein; 433 AA.

XX AC AAW95698;

XX DT 16-JUN-1999 (first entry)

XX DE *Bacillus* sp. Lion Y protease.

XX KW PD498; subtilisin; Lion Y; protease; skin-care; feed; additive; soap;

KW

KW

KW

XX

OS

XX

PN

XX

PD

XX

PP

XX

XX

PR

XX

PA

XX

PI

XX

XX

DR

XX

PT

XX

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

cosmetic; hair dye; sunscreen; acne; antiperspirants; insect repellent;
 deodorant; detergent; food; breadmaking; textile-treating; oral; dermal;
 pharmaceutical; agrochemical.

Bacillus sp.

WO9900489-A1.

07-JAN-1999.

22-JUN-1998; 98WO-DK000270.

25-JUN-1997; 97DK-00000753.

07-JUL-1997; 97US-0051830P.

(NOVO) NOVO-NORDISK AS.

Olsen AA, Fatum TM, Deussen H, Roggen EL;

WPI; 1999-095735/08.

New modified polypeptide with attached low molecular weight polymer - has
 reduced respiratory allergenicity, useful in skin care products,
 detergents, as food additives or textile-treating compositions.

Claim 10; Page 48-49; 60pp; English.

The sequence is that of Lion Y protease. This can be used as an active
 ingredient: (i) in personal care products (especially skin-care products,
 such as soaps, cosmetics, hair dyes, sunscreens, anti-acne products,
 antiperspirants, insect repellants or deodorants); (ii) in detergents (as
 laundry, dishwashing or hard-surface cleaners); (iii) food or feed
 additives (e.g. for breadmaking); (iv) in textile-treating compositions,
 or (v) in oral or dermal pharmaceuticals and agrochemicals

Sequence 433 AA;

Query Match 99.4%; Score 2247; DB 2; Length 433;

Best Local Similarity 99.3%; Pred. No. 5.7e-168;
 Matches 430; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIKADVAAQNNYGLYGGQGVAVADTGLTGRNDSMHEAFRGKITALYALGRTN 60

DB 1 NDVARGIKADVAAQNNYGLYGGQGVAVADTGLTGRNDSMHEAFRGKITALYALGRTN 60

QY 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGPSNLNLFQAWNA 120

DB 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGPSNLNLFQAWNA 120

QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVNNDMTTLFAAGNEGPNSTGTSAPGTAKNAIT 180

DB 121 GARIHTNSWGAPVNGAYTANSRQVDEYVNNDMTTLFAAGNEGPNSTGTSAPGTAKNAIT 180

QY 181 VGATENYRPSFGSIADNPNHIAQSSRGATRDGRIPKDVTPAGTFTLSARSLAPDSSF 240

DB 181 VGATENYRPSFGSIADNPNHIAQSSRGATRDGRIPKDVTPAGTFTLSARSLAPDSSF 240

QY 241 ANYNSKYAVMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300

DB 241 ANYNSKYAVMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300

QY 301 NGDQGWGRVTLNKSINVAVNEATATGCKATYSFQAQKPLKISLVWTDAPGSTTAS 360

DB 301 NGDQGWGRVTLNKSINVAVNEATATGCKATYSFQAQKPLKISLVWTDAPGSTTAS 360

QY 361 YTLVNDLDELVITAPNGQKYGNDFSYPYDNNWGRNNVENVFINAPQSGTYIIIEVQAYNV 420

DB 361 YTLVNDLDELVITAPNGQKYGNDFSYPYDNNWGRNNVENVFINAPQSGTYIIIEVQAYNV 420

QY 421 PSGPQRFSLAIVH 433

DB 421 PSGPQRFSLAIVH 433

RESULT 5
 AAY69207
 ID AAY69207 standard; protein; 433 AA.
 XX
 AC AAY69207;
 DT 30-MAY-2000 (first entry)
 XX
 DE Amino acid sequence of protease Lion Y.
 XX
 KW Protease Lion Y; polypeptide-polymer conjugate; washing performance;
 KW respiratory allergenicity; allergic reaction; detergent formulation;
 KW laundry; dishwashing; hard surface cleaner; agricultural chemical;
 KW skin care; cosmetic; oral pharmaceutical; dental pharmaceutical;
 KW textile processing.
 XX
 OS Bacillus sp.
 XX
 XX WO200004138-A1.
 XX
 PD 27-JAN-2000.
 XX
 PF 16-JUL-1999; 99WO-DK000405.
 XX
 PR 17-JUL-1998; 98DK-00000951.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Bauditz P, Fatum TM, Olsen AA, Deussen H, Petersen DA;
 XX
 XX WPI; 2000-195024/17.
 XX
 XX New polypeptide-polymer conjugate, particularly enzyme conjugate, useful
 PT in detergent formulations.
 PT
 XX Disclosure; Page 74-76; 79pp; English.
 XX
 CC The present sequence represents a protease Lion Y protein. The protein
 CC may be used to produce the conjugates of the invention. The specification
 CC describes polypeptide-polymer conjugates which have improved washing
 CC performance and reduced respiratory allergenicity compared with the
 CC unconjugated polypeptide. The polymer provides, in water, a conformation
 CC that shields the molecular surface effectively, preventing association of
 CC antibodies that can induce an allergic reaction. The conjugates are used
 CC in industrial compositions, particularly detergent formulations (laundry,
 CC dishwashing or hard surface cleaners), but also in agricultural
 CC chemicals, skin care products (cosmetics and toiletries), oral and dental
 CC pharmaceuticals, or textile processing and treatment compositions
 XX
 SQ Sequence 433 AA;
 Query Match 99.4%; Score 2247; DB 3; Length 433;
 Best Local Similarity 99.3%; Pred. No. 5.7e-168;
 Matches 430; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAQNNGYLGQGVAVADTGLDTRNDSSMHEAFRGKITAYALGRTN 60
 DB 1 NDVARGIVKADVAQNNGYLGQGVAVADTGLDTRNDSSMHEAFRGKITAYALGRTN 60
 QY 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120
 DB 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120
 QY 121 GARIHNSGAPVNGAYTANSRQVDEYVRNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
 DB 121 GARIHNSGAPVNGAYTANSRQVDEYVRNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
 QY 181 VGATENYRSPFGSIADPNPHIAQFSSRGATRGRIKPDVTAPCTFTLSARSSLAPDSGF 240
 DB 181 VGATENYRSPFGSIADPNPHIAQFSSRGATRGRIKPDVTAPCTFTLSARSSLAPDSGF 240
 QY 241 ANTNSKYAYMGTSMTPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVLGY 300

241 ANYNSKYAYMGTSMTPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVLGY 300
 301 NGDQGWGRVTLNKSINVAVNEATATATGOKATYSFQAQAGKPLKISLVWTDAPGSTTAS 360
 301 SGDQGWGRVTLNKSINVAVNEATATATGOKATYSFQAQAGKPLKISLVWTDAPGSTTAS 360
 361 YTLVNDLDELITAPNGQKYVGNDFSYVDNNDWGRNNVENVFINAPQSGTYIIIEVQAYNV 420
 361 YTLVNDLDELITAPNGQKYVGNDFSYVDNNDWGRNNVENVFINAPQSGTYIIIEVQAYNV 420
 421 PSQPQRFSLAIVH 433
 421 PSQPQRFSLAIVH 433
 RESULT 6
 AAY44619
 ID AAY44619 standard; protein; 433 AA.
 XX
 AC AAY44619;
 DT 07-APR-2000 (first entry)
 XX
 DE Bacillus Lion Y enzyme.
 XX
 KW Lion Y enzyme; protease; allergic response; industrial composition;
 KW co-polymer; ethylene oxide; EO; propylene oxide; PO; conjugate;
 KW allergenicity; detergent; cosmetic; toiletries; textile treatment;
 KW agrochemical; pharmaceutical; food; feed additive.
 XX
 OS Bacillus sp.
 XX
 XX WO9967370-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 23-JUN-1999; 99WO-DK000359.
 XX
 PR 23-JUN-1998; 98DK-00000809.
 XX
 PA (NOVO) NOVO-NORDISK AS.
 XX
 PI Deussen H, Olsen AA, Fatum TM, Roggen EL;
 XX
 XX WPI; 2000-136981/12.
 PT New conjugate of polypeptide, especially an enzyme, with copolymer of
 PT ethylene oxide and propylene oxide, used in e.g. cleaning compositions,
 PT has reduced allergenicity.
 XX
 PS Claim 10; Page 56-57; 62pp; English.
 XX
 CC The present sequence is a Bacillus Lion Y enzyme, which is a protease
 CC capable of inducing an allergic response upon inhalation. The enzyme can
 CC be covalently coupled to a co-polymer comprising ethylene oxide (EO) and
 CC propylene oxide (PO) to reduce its allergenicity. This enzyme-polymer
 CC conjugate can be used in industrial compositions such as detergents,
 CC cosmetics, toiletries, textile treatment compositions, agrochemicals,
 CC oral and dermal pharmaceuticals and food and feed additives
 XX
 SQ Sequence 433 AA;
 Query Match 99.4%; Score 2247; DB 3; Length 433;
 Best Local Similarity 99.3%; Pred. No. 5.7e-168;
 Matches 430; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAQNNGYLGQGVAVADTGLDTRNDSSMHEAFRGKITAYALGRTN 60
 DB 1 NDVARGIVKADVAQNNGYLGQGVAVADTGLDTRNDSSMHEAFRGKITAYALGRTN 60
 QY 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120

Db 61 NASDPNGCHTHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
 QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLPFAAGNEGPNSTISAPGTAKNAIT 180
 Db 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLPFAAGNEGPNSTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIPKDVTPAGTFFILSARSSLPDSSF 240
 Db 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIPKDVTPAGTFFILSARSSLPDSSF 240
 QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFIPKRGITPKPSLIKAALIAGATDVGLGYP 300
 Db 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFIPKRGITPKPSLIKAALIAGATDVGLGYP 300
 QY 301 NGDQGWGRVTLNKSINVAIVNEATATATGOKATYSFOAQAGKPLKISLWTDAPGSTTAS 360
 Db 301 NGDQGWGRVTLNKSINVAIVNEATATATGOKATYSFOAQAGKPLKISLWTDAPGSTTAS 360
 QY 361 YTLVNDLVLITAPNGQKYGNDPFPYDNNWNGRNNVNFINAPQSGTYIIIEVQAYNV 420
 Db 361 YTLVNDLVLITAPNGQKYGNDPFPYDNNWNGRNNVNFINAPQSGTYIIIEVQAYNV 420
 QY 421 PSGQRFSLAIVH 433
 Db 421 PSGQRFSLAIVH 433

RESULT 7

AAM50084

ID AAM50084 standard; protein; 433 AA.

XX

AC AAM50084;

XX 12-AUG-2002 (first entry)

DT

DE Bacillus sp SD-521 (FERM BP-11162) alkaline protease protein fragment.

XX

KW Alkaline protease; detergent; laundry; bleaching; dishwasher.

XX

OS Bacillus sp.

XX

PN BP1209233-A2.

XX

PD 29-MAY-2002.

XX

XX 22-NOV-2001; 2001EP-00127851.

XX

PR 22-NOV-2000; 2000JP-00355166.

XX

PR 12-APR-2001; 2001JP-00114048.

XX

XX (KAOS) KAO CORP.

XX

XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N.

PI Okuda M, Saeki K;

XX

XX WPI; 2002-437518/47.

XX

XX New modified alkaline proteases useful in detergent compositions.

PT

PS Claim 5; Page 16-18; 25pp; English.

XX

XX This invention describes novel Bacillus sp. alkaline proteases useful in

CC detergent compositions, especially in laundry, bleaching or automatic

CC dishwasher detergents. The novel proteases have an increased detergency *

CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This

CC sequence represents a fragment of the alkaline protease SD-521 from

CC Bacillus sp strain SD-521 (FERM BP-11162) described in the method of the

CC invention

XX

SQ Sequence 433 AA;

XX

Query Match 99.1%; Score 2240; DB 5; Length 433;

Best Local Similarity 98.8%; Pred. No. 2e-167;

XX

XX

XX

XX

XX

XX

XX

XX

XX

Matches 428; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVADVAQNNYGLYGQGVVAVADTGLDGRNDSSMHEAPRGKITALYALGRTN 60
 Db 1 NDVARGIVKADVADVAQNNYGLYGQGVVAVADTGLDGRNDSSMHEAPRGKITALYALGRTN 60
 QY 61 NASDPNGCHTHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
 Db 61 NASDPNGCHTHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
 QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLPFAAGNEGPNSTISAPGTAKNAIT 180
 Db 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLPFAAGNEGPNSTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIPKDVTPAGTFFILSARSSLPDSSF 240
 Db 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIPKDVTPAGTFFILSARSSLPDSSF 240
 QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFIPKRGITPKPSLIKAALIAGATDVGLGYP 300
 Db 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFIPKRGITPKPSLIKAALIAGATDVGLGYP 300
 QY 301 NGDQGWGRVTLNKSINVAIVNEATATATGOKATYSFOAQAGKPLKISLWTDAPGSTTAS 360
 Db 301 NGDQGWGRVTLNKSINVAIVNEATATATGOKATYSFOAQAGKPLKISLWTDAPGSTTAS 360
 QY 361 YTLVNDLVLITAPNGQKYGNDPFPYDNNWNGRNNVNFINAPQSGTYIIIEVQAYNV 420
 Db 361 YTLVNDLVLITAPNGQKYGNDPFPYDNNWNGRNNVNFINAPQSGTYIIIEVQAYNV 420
 QY 421 PSGQRFSLAIVH 433
 Db 421 PSGQRFSLAIVH 433

RESULT 8

AAM89548

ID AAM89548 standard; protein; 636 AA.

XX

AC AAM89548;

XX

XX 12-APR-1999 (first entry)

DT

DE Bacillus sp. alkaline protease Y.

XX

KW Alkaline protease Y; detergent; surfactant; leather processing;

KW debittering; flavour.

XX

OS Bacillus sp.

XX

XX WO9856927-A2.

XX

XX 17-DEC-1998.

PD

XX

XX 09-JUN-1998; 98WO-US012005.

XX

XX 12-JUN-1997; 97US-00873479.

XX

XX (NOVO) NOVO NORDISK BIOTECH INC.

XX

XX Sioma A, Christianson L;

XX

XX WPI; 1999-080908/07.

DR

XX

XX Novel protease from Bacillus subtilis LC20 - useful in laundry and

PT dishwashing detergents and for leather processing.

XX

XX Claim 3; Page 55-56; 77pp; English.

XX

XX This is the amino acid sequence of a Bacillus sp. alkaline protease Y

XX that is said to have good alkali and surfactant resistance and improved

CC detergency. It shows 77% identity to a newly isolated protease (see

CC AAM89547) of Bacillus sp. JPI70 (NCIB 12513). The invention provides

CC

CC vectors, recombinant host cells and methods for the recombinant
 CC production of such proteases. The protease are used in laundry and
 CC dishwashing detergents, for institutional and industrial cleaning, and
 CC for leather processing, as well as for debittering and enhancing the
 CC degree of hydrolysis of protein hydrolysates, for flavour development
 CC through hydrolysis of proteins, degradation of undesired peptides and
 CC enzymatic synthesis of peptides. They have enhanced stability towards
 CC oxidation under alkaline conditions, e.g. towards bleaching agents of the
 CC peroxy type. The invention also provides mutant cells in which the
 CC protease activity is diminished. Such cells can be used for the
 CC production of heterologous recombinant proteins
 XX
 SQ Sequence 636 AA;

Query Match 99.0%; Score 2238; DB 2; Length 636;
 Best Local Similarity 99.1%; Pred. No. 5e-167;
 Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAQNNGYLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 DB 204 NDVARGIVKADVAQNNGYLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 263
 QY 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNLFPSQAWNA 120
 DB 264 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNLFPSQAWNA 323
 QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTISAPGTAKNAIT 180
 DB 324 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTISAPGTAKNAIT 383
 QY 181 VGATENYRPSFGSIADPNHIAQFSSRGATRGRIKPDVTAPGTFTLSARSLAPDSSFW 240
 DB 384 VGATENYRPSFGSIADPNHIAQFSSRGATRGRIKPDVTAPGTFTLSARSLAPDSSFW 443
 QY 241 ANYNSKYAYMGTSMATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGYP 300
 DB 444 ANYNSKYAYMGTSMATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGYP 503
 QY 301 NGDQGWGRVTLNKSINVAVYNEATATGOKATYSFOAAGKPLKISLVWTDAPSGTTAS 360
 DB 504 SGDQGWGRVTLNKSINVAVYNEATATGOKATYSFOAAGKPLKISLVWTDAPSGTTAS 563
 QY 361 YTLVNDLDELVTAPNGQKVGNDFSYFYDNNNDGRNNVENVFINAPQSGTYIIIEVQAYNV 420
 DB 564 YTLVNDLDELVTAPNGQKVGNDFSYFYDNNNDGRNNVENVFINAPQSGTYIIIEVQAYNV 623
 QY 421 PSQPQRFSIAIVH 433
 DB 624 PSQPQRFSIAIVH 636

RESULT 9
 AAM50082
 ID AAM50082 standard; protein; 433 AA.
 AC
 AC AAM50082;
 XX

XX 12-AUG-2002 (first entry)
 XX Bacillus sp D6-(FERM P1592) alkaline protease protein fragment.
 XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX Bacillus sp.

XX EP1209233-A2.
 XX 29-MAY-2002.
 XX 22-NOV-2001; 2001EP-00127851.
 XX 22-NOV-2000; 2000JP-00355166.
 PR 12-APR-2001; 2001JP-00114048.

XX (KAOS) KAO CORP.
 XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 XX Okuda M, Saeki K;
 XX WPI; 2002-437518/47.
 XX New modified alkaline proteases useful in detergent compositions.
 XX Claim 5; Page 13-15; 25pp; English.
 XX This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency &
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease E-1 from Bacillus
 CC sp strain D6-(FERM-P1592) described in the method of the invention
 XX
 SQ Sequence 433 AA;

Query Match 98.8%; Score 2234; DB 5; Length 433;
 Best Local Similarity 98.6%; Pred. No. 6e-167;
 Matches 427; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAQNNGYLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAQNNGYLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNLFPSQAWNA 120
 DB 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNLFPSQAWNA 120
 QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTISAPGTAKNAIT 180
 DB 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSIADPNHIAQFSSRGATRGRIKPDVTAPGTFTLSARSLAPDSSFW 240
 DB 181 VGATENYRPSFGSIADPNHIAQFSSRGATRGRIKPDVTAPGTFTLSARSLAPDSSFW 240
 QY 241 ANYNSKYAYMGTSMATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGYP 300
 DB 241 ANYNSKYAYMGTSMATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGYP 300
 QY 301 NGDQGWGRVTLNKSINVAVYNEATATGOKATYSFOAAGKPLKISLVWTDAPSGTTAS 360
 DB 301 SGDQGWGRVTLNKSINVAVYNEATATGOKATYSFOAAGKPLKISLVWTDAPSGTTAS 360
 QY 361 YTLVNDLDELVTAPNGQKVGNDFSYFYDNNNDGRNNVENVFINAPQSGTYIIIEVQAYNV 420
 DB 361 YTLVNDLDELVTAPNGQKVGNDFSYFYDNNNDGRNNVENVFINAPQSGTYIIIEVQAYNV 420
 QY 421 PSQPQRFSIAIVH 433
 DB 421 PSQPQRFSIAIVH 433

RESULT 10
 AAM50086
 ID AAM50086 standard; protein; 433 AA.
 XX
 AC AAM50086;
 XX

XX 12-AUG-2002 (first entry)
 XX Bacillus sp alkaline protease protein A-2 fragment.
 XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX Bacillus sp.
 XX EP1209233-A2.

| | |
|----|---|
| DE | Bacillus JP170 protease. |
| XX | |
| XX | Protease; detergent; surfactant; leather processing; debittering; |
| KW | flavour. |
| XX | |
| OS | Bacillus sp. |
| XX | |
| FH | Key Location/Qualifiers |
| FT | Peptide 1..33 |
| FT | /note= "signal peptide" |
| FT | Region 34..208 |
| FT | /note= "prepro region" |
| FT | Protein 209..641 |
| FT | /note= "mature protein" |
| XX | |
| PN | WC9856927-A2. |
| XX | |
| PD | 17-DEC-1998. |
| XX | |
| PF | 09-JUN-1998; 98WO-US012005. |
| XX | |
| PR | 12-JUN-1997; 97US-00873479. |
| XX | |
| PA | (NOVO) NOVO NORDISK BIOTECH INC. |
| XX | |
| PI | Sloma A, Christianson L; |
| XX | |
| DR | WIPI: 1999-080908/07. |
| DR | N-PSDB; AAV82382. |
| XX | |
| PT | Novel protease from Bacillus subtilis LC20 - useful in laundry and |
| PT | dishwashing detergents and for leather processing. |
| XX | |
| PS | Claim 7; Page 53-54; 77pp; English. |
| CC | This is the amino acid sequence of a novel protease of Bacillus sp. JP170 |
| CC | (NCBI 12513), as deduced from the nucleotide sequence of an isolated gene |
| CC | (see AAV82382). The entire protein, including the signal peptides and |
| CC | prepro region, has 77% identity to alkaline protease Y (see AAM89548) |
| CC | from Bacillus. The invention provides vectors, recombinant host cells and |
| CC | methods for the recombinant production of the protease. The protease is |
| CC | used in laundry and dishwashing detergents, for institutional and |
| CC | industrial cleaning, and for leather processing, as well as for |
| CC | debittering and enhancing the degree of hydrolysis of protein |
| CC | hydrolysates, for flavour development through hydrolysis of proteins. |
| CC | Degradation of undesired peptides and in enzymatic synthesis of peptides. |
| CC | It has enhanced stability towards oxidation under alkaline conditions, |
| CC | e.g. towards bleaching agents of the peroxy type. The invention also |
| CC | provides mutant cells in which the protease activity is diminished. Such |
| CC | cells can be used for the production of heterologous recombinant proteins |
| XX | |
| SQ | Sequence 641 AA; |
| | Query Match 90.0%; Score 2033; DB 2; Length 641; |
| | Best Local Similarity 88.7%; Pred. No. 6.5e-151; |
| | Matches 384; Conservative 25; Mismatches 24; Indels 0; Gaps 0 |
| Qy | 1 NDVARGIVKADVAAQNNYGLYGGGVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 60 |
| | ::::::::::::::::::::: |
| Db | 209 NDVARGIVKADVAAQNNGFLYGGGVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 268 |
| | ::::::::::::::::::::: |
| Qy | 61 NASDPNGHGTHVAGSVLGNALKMGWAQANLVFQSIMDSGGGLGGLPNSNLTLFSQAANA 120 |
| | ::::::::::::::::::::: |
| Db | 269 NANDPNGHGTHVAGSVLGNATKMGWAQANLVFQSIMDSGGGLGGLPANLTQLFSQAISA 328 |
| | ::::::::::::::::::::: |
| Qy | 121 GARHTNSWGAPVNGAYTANGSRQVDEVYRNDMTVLFAAGNEPGSGTTISAPGTAKNAIT 180 |
| | ::::::::::::::::::::: |
| Db | 329 GARHTNSWGAPVNGAYTTDSRNVDYTRKNDMTILFAAGNEPGSGTTISAPGTAKNAIT 388 |
| | ::::::::::::::::::::: |
| Qy | 181 VGATENYRSPFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFILSARSSLAPDSSF 240 |
| | ::::::::::::::::::::: |
| Db | 389 VGateMLRFSGSYADNIHNHVAQFSSRGPTFDRIKPDVWAPGTIYLSAESSLAPDSSF 448 |

| | | | |
|-----|----|--|-----|
| 241 | QY | ANYNSKYAWMGTSMTPTIVAGVNAQLREHFTPKRGITPKPSLLIKAALIAGADVGLGPP | 300 |
| 449 | Db | ANHDSKYAWMGTSMTPTIVAGVNAQLREHFTVQRGVTPKPSLLIKAALIAGAADVGLGPP | 508 |
| 301 | QY | NGDQGRVYTLANKSLNVAVNEATATATGOKATYSFQAQAGKPKLSLVWTDAPGGTTAS | 360 |
| 509 | Db | NGNQGRVYTLDKSLNVAFPVNETSPSTYSQKATYSFQAQAGKPKLSLVWTDAPGGTTAS | 568 |
| 361 | QY | YTLVNDLDELVIITAPNGQKTVGNDFSYPYDNNWDRNNVNFVINAPOSQTYIIEVQAYNV | 420 |
| 569 | Db | LTLVNDLDELVIITAPNGTKYVGNDFPTAYDNNWDRNNVNFVINAPOSQTYIIEVQAYNV | 628 |
| 421 | QY | PSGPQSFSLAIYH 433 | |
| 629 | Db | PVSPQTFSLAIYH 641 | |

| | |
|-----------|---|
| RESULT 12 | |
| AAAM50085 | |
| ID | AAAM50085 standard; protein; 434 AA. |
| XX | |
| XX | AAAM50085; |
| XX | |
| XX | 12-AUG-2002 (first entry) |
| DT | |
| XX | |
| XX | Bacillus sp alkaline protease protein A-1 fragment. |
| DE | |
| XX | |
| XX | Alkaline protease; detergent; laundry; bleaching; dishwasher. |
| KW | |
| XX | |
| XX | Bacillus sp. |
| OS | |
| XX | |
| XX | EP1209233-A2. |
| PN | |
| XX | |
| XX | 29-MAY-2002. |
| PD | |
| XX | |
| XX | 22-NOV-2001; 2001EP-00127851. |
| XX | |
| XX | |
| PR | 22-NOV-2000; 2000JP-00355166. |
| PR | 12-APR-2001; 2001JP-00114048. |
| XX | |
| XX | (KAOS) KAO CORP. |
| FA | |
| XX | |
| PI | Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N; |
| PI | Okuda M, Saeki K; |
| PI | |
| XX | |
| DE | WPI; 2002-437518/47. |
| XX | |
| XX | New modified alkaline proteases useful in detergent compositions. |
| ET | |
| XX | |
| PS | Claim 5; Page 18-19; 25pp; English. |
| XX | |
| XX | This invention describes novel Bacillus sp. alkaline proteases useful in detergent compositions, especially in laundry, bleaching or automatic dishwasher detergents. The novel proteases have an increased detergency (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This sequence represents a fragment of the alkaline protease A-1 from Bacillus sp NCIB12289 described in the method of the invention |
| CC | |
| CC | |
| CC | |
| CC | |
| CC | |
| CC | |
| CC | |
| CC | |
| XX | |
| XX | Sequence 434 AA; |
| XX | |

| Query Match | 88.7%; | Score 2005.5; | DB 5; | Length 434; |
|-----------------------|------------------|---|-----------|-------------|
| Best Local Similarity | 87.6%; | Pred. No. 5.5e-149; | | |
| Matches 380; | Conservative 31; | Mismatches 22; | Indels 1; | Gaps 1 |
| Qy | 1 | NDVARGIVKADVAQNNGYGLYGGQVAVADTGLDTGRNDSMHEAFRGKITAIYALGRTN | 60 | |
| Db | 1 | NDVARGIVKADVAQSYGLYGGQVAVADTGLDTGRNDSMHEAFRGKITAIYALGRTN | 60 | |
| Qy | 61 | NASDPNGHGTTHVAGSVLGNAL-NKGMAPOANLVFQSTWDSGGGLGGLPSNLATLFSQAWN | 119 | |
| Db | 61 | NANDPNGHGTTHVAGSVLGNGTNKGMAPOANLVFQSVMDNNGGLGGVPSNVTLFSQAYS | 120 | |
| Qy | 120 | AGARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNENGPNSGTISAPGTAKNAI | 179 | |

| | | |
|-----|---|------|
| 121 | AGARHTNSWGPVNGAYTTDSRVVDYVRKNDNAVLFPAAGNEGPGGTISAFETAKNAI | 180 |
| Db | : | |
| 180 | TVGATENRPFSGSIADNPNEHIAQFSRGATDRGRIKPDVYAPGCTFILSARSLAPDSSF | 2399 |
| Qy | : | |
| 181 | TVGATENRPFSGSYADINHVAFSRGPTKDRGRIKPDVYAPGCTFILSARSLAPDSSF | 2400 |
| Db | : | |
| 240 | WANTYSKAYMGGTSMATPIVAGNVAQLREHFIKRGGITPKPLIKAALAGATDVGLGY | 2999 |
| Qy | : | |
| 241 | WANHDSKAYMGGTSMATPIVAGNVAQLREHFIKRGGITPKPLILKAALAGATDILGY | 3000 |
| Db | : | |
| 300 | PNGDQGWGRVTLNKSINVAYNVEATALATGQATYSPQAGAKPLKISLWTDAPGSTTA | 3599 |
| Qy | : | |
| 301 | PSGNQGWGRVTLDKSLNVAFVNETSSLSNQATYSPAQSGRPLKISLWSDAPASTSA | 3600 |
| Db | : | |
| 360 | SVTLVNDLDLVITAPNGOKYVGNDFSPYDNNWDGRNNVENFINAPQSGTYLIEVQAYN | 419 |
| Qy | : | |
| 361 | SVTLVNDLDLVITAPNGTKYVGNDFTPAYDNNWDGRNNVENFINAPQSGTYTVEVQAYN | 4200 |
| Db | : | |
| 420 | VPSGQPRFSLAIHV 433 | |
| Qy | : | |
| 421 | VFOGPOAFSLAIYN 434 | |
| Db | : | |

RESULT 13
AAM50081
ID AAM50081 standard; protein; 434 AA.
XX
XX AAM50081;
XX
XX
DT 12-AUG-2002 (first entry)
XX
XX Bacillus sp KSM-KP9860 alkaline protease protein fragment.
XX
XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
KW
XX
XX Bacillus sp.
OS
XX
XX EF1209233-A2.
PN
XX
XX 29-MAY-2002.
PD
XX
XX 22-NOV-2001; 2001EP-00127851.
PF
XX
XX 22-NOV-2000; 2000JP-00355166.
PR
XX
XX 12-APR-2001; 2001JP-00114048.
PR
XX
XX (KAOS) KAO CORP.
PA
XX
XX Harada Y, Ogawa A, Kageyama Y, Sato T, Azaki H, Sumitomo N;
PI
PI Okuda M, Saeiki K;
PI
XX
XX
DR WPI; 2002-437518/47.
XX
XX New modified alkaline proteases useful in detergent compositions.
FT
XX
XX Claim 5; Page 12-13; 25pp; English.
PS

```

XX CC This invention describes novel Bacillus sp. alkaline proteases useful in
XX CC detergent compositions, especially in laundry, bleaching or automatic
XX CC dishwasher detergents. The novel proteases have an increased detergency &
XX CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
XX CC sequence represents a fragment of the alkaline protease KP9860 from
XX CC Bacillus sp strain KSM-KP9860 described in the method of the invention
XX SQ Sequence 434 AA;

Query Match      88.5%; Score 1999, 5; DB 5; Length 434;
Best Local Similarity 88.0%; Pred. No. 1.6e-148;
Matches 382; Conservative 24; Mismatches 24; Indels 1; Gaps 1

Qy 1 NDVARGIVKADVAQNNTGLYGGQGVVAVDLTGTGRNDSMMHFAFRGKITALYALGRTN 60

```

Db 1 NDVARGIVKADVQAQSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALVALGRTN 60
 QY 61 NASDPNGHGHVAGSVLGN-ALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSAQWN 119
 Db 61 NANTNGHGHVAGSVLGNATNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSAQFS 120
 QY 120 AGARIHTNSWAPNGAYTANSRQVDEYVRNDMTVLPAAAGNEGPNSTISAPGTAKNAI 179
 Db 121 AGARIHTNSWGAANGAYTTDSRNVDYVRKNDMTILFAAGNEGPNSTISAPGTAKNAI 180
 QY 180 TVGATENYRPFSGSIADPNHIAQSFSSRGATRDGRIKPDVTPAGTFILSARSSLAPDSSF 239
 Db 181 TVGATENLRFPSGSIADPNHIAQSFSSRGATRDGRIKPDVTPAGTFILSARSSLAPDSSF 240
 QY 240 WANYNSKYAYMGGTSMATPIVAGNVAQLREHFVKRGTTPKPSLIIKALIIAGATDVLGY 299
 Db 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKRGTTPKPSLIIKALIIAGADVLGY 300
 QY 300 PNGDQGWGRVTLNKLNVAYVNEATATGOKATYSFOAQGKPLKISLVWTDAPGSTTA 359
 Db 301 PNGDQGWGRVTLNKLNVAYVNEATATGOKATYSFOAQGKPLKISLVWTDAPGSTTA 360
 QY 360 SVTLVNDLVLITAPNGQKYGNDPSYPYDNNWGDGNNVNFVINAPOSQGTIIIEVQAYN 419
 Db 361 SVTLVNDLVLITAPNGQKYGNDPSYPYDNNWGDGNNVNFVINAPOSQGTIIIEVQAYN 420
 QY 420 VPSGQPSLAIVH 433
 Db 421 VPVGPQNFSLAIVN 434

RESULT 14

AAV17089
 ID AAY17089 standard; protein; 639 AA.

XX AC AAY17089;
 XX AC AAY17089;
 DT 20-MAR-2003 (revised)
 DT 21-JUL-1999 (first entry)
 XX DE Bacillus alkaline protease.
 XX KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
 XX KW washing composition; oxidising agent.
 XX OS Bacillus sp.
 XX PN WO9918218-A1.
 XX PD 15-APR-1999.
 XX PF 07-OCT-1998; 98WO-JP004528.
 XX PR 07-OCT-1997; 97JP-00274570.
 XX PA (KAOS) KAO CORP.
 XX PI Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
 XX PI Shikata S, Nomura M;
 XX DR WPI; 1999-287736/27.
 XX DR N-PSDB; AAX37277.

XX PT Alkali protease from Bacillus used in washing powders.

XX PS Disclosure; Page 53-58; 71pp; Japanese.

XX CC The invention relates to alkaline proteases produced by strains of
 CC Bacillus. The proteases ability to digest casein is not inhibited by
 CC oleic acid and they have a high stability to oxidising agents. The
 CC alkaline protease of the invention has the following properties: (a) it
 CC is active over the pH range 4-13 and has at least 80% of its optimum
 CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is

CC stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d)
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has
 CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
 CC used as enzymes in washing compositions for use in automatic dishwashers
 CC and for washing clothes. The stability to oxidising agents allows the
 CC enzyme to be an effective component of washing compositions including
 CC bleaches. The present sequence represents an alkaline protease. (Updated
 CC on 20-MAR-2003 to correct DR field.)

XX SQ Sequence 639 AA;

Query Match 88.1%; Score 1991.5; DB 2; Length 639;
 Best Local Similarity 87.8%; Pred. No. 1.2e-147;
 Matches 391; Conservative 27; Mismatches 25; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVQAQSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALVALGRTN 60
 Db 206 NDVARGIVKADVQAQSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALVALGRTN 265
 QY 61 NASDPNGHGHVAGSVLGN-ALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSAQWN 119
 Db 266 NANTNGHGHVAGSVLGNATNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSAQFS 325
 QY 120 AGARIHTNSWAPNGAYTANSRQVDEYVRNDMTVLPAAAGNEGPNSTISAPGTAKNAI 179
 Db 326 AGARIHTNSWGAANGAYTTDSRNVDYVRKNDMTILFAAGNEGPNSTISAPGTAKNAI 385
 QY 180 TVGATENYRPFSGSIADPNHIAQSFSSRGATRDGRIKPDVTPAGTFILSARSSLAPDSSF 239
 Db 386 TVGATENLRFPSGSIADPNHIAQSFSSRGATRDGRIKPDVTPAGTFILSARSSLAPDSSF 445
 QY 240 WANYNSKYAYMGGTSMATPIVAGNVAQLREHFVKRGTTPKPSLIIKALIIAGATDVLGY 299
 Db 446 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFVKRGTTPKPSLIIKALIIAGADVLGY 505
 QY 300 PNGDQGWGRVTLNKLNVAYVNEATATGOKATYSFOAQGKPLKISLVWTDAPGSTTA 359
 Db 506 PNGDQGWGRVTLNKLNVAYVNEATATGOKATYSFOAQGKPLKISLVWTDAPGSTTA 565
 QY 360 SVTLVNDLVLITAPNGQKYGNDPSYPYDNNWGDGNNVNFVINAPOSQGTIIIEVQAYN 419
 Db 566 SVTLVNDLVLITAPNGQKYGNDPSYPYDNNWGDGNNVNFVINAPOSQGTIIIEVQAYN 625
 QY 420 VPSGQPSLAIVH 433
 Db 626 VPVGPQNFSLAIVN 639

RESULT 15

AAV17091
 ID AAY17091 standard; protein; 640 AA.

XX AC AAY17091;
 XX AC AAY17091;
 DT 20-MAR-2003 (revised)
 DT 21-JUL-1999 (first entry)
 XX DE Bacillus alkaline protease.
 XX KW Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
 XX KW washing composition; oxidising agent.
 XX OS Bacillus sp.
 XX PN WO9918218-A1.
 XX PD 15-APR-1999.
 XX PF 07-OCT-1998; 98WO-JP004528.
 XX PR 07-OCT-1997; 97JP-00274570.
 XX PA (KAOS) KAO CORP.

```
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;
XX WPI: 1999-287736/27.
DR N-PSDB; AAX37279.
XX
FT Alkali protease from Bacillus used in washing powders.
XX
PS Disclosure; Page 63-68; 71pp; Japanese.
XX
XX The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidising agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidising agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease. (Updated
CC on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 640 AA;

Query Match      88.0%; Score 1988.5; DB 2; Length 640;
Best Local Similarity 87.3%; Pred. No. 2e-147;
Matches 379; Conservative 29; Mismatches 25; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVQANNVGLYGGQVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
DB 207 NDVARGIVKADVQANNVGLYGGQVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 266
QY 61 NASDPNGHGHVAGSVLGN-ALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWN 119
DB 267 NANTDNGHGHVAGSVLGNSTNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAYS 326
QY 120 AGARIHTNSGAPVNGAYTANSQVDEYVRNNDMTVLFAGNEGPNSTISAPGTAKNAI 179
DB 327 AGARIHTNSGAAVNGAYTTDSRNDYVRKNDMTILFAGNEGPNSTISAPGTAKNAI 386
QY 180 TVGATENYRPSFGSIADNPNHIAQFSRGATRDGRIPDVTAPGTFILSARSSLAPDSF 239
DB 387 TVGATENLRPSFGSIADNPNHIAQFSRGATRDGRIPDVTAPGTFILSARSSLAPDSF 446
QY 240 WANYNSKYAYMGCTSMATPIVAGNVAQLREHFTKRGITKPSLIKAALIAGATDVGLGY 299
DB 447 WANHDSKYAYMGCTSMATPIVAGNVAQLREHFTKRGITKPSLIKAALIAGATDVGLGY 506
QY 300 PNGDQGHGVRTLNKSLNVAAYNEATATGQKATYSFQAQAGKPLKISLVWTDAPGSTTA 359
DB 507 PNGNQGHGVRTLDKSLNVAAYNESSTLSQKATYSFTATAGKPLKISLVWSDAPASTTA 566
QY 360 SYTLVNDLDLVTAPNCKVGVNDYFYPYDNNWGDGRNNVENFVINAPOSQGTIIIEVQAYN 419
DB 567 SVTLVNDLDLVTAPNCKVGVNDYFYPYDNNWGDGRNNVENFVINAPOSQGTIIIEVQAYN 626
QY 420 VPSGPQRFSIAIVH 433
DB 627 VPVGQNFSLAIVN 640
```

Search completed: March 31, 2004, 16:04:31
Job time : 48.1304 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:59:39 ; Search time 13.982 Seconds
(without alignments)
1597.947 Million cell updates/sec

Title: US-09-985-689A-4

Perfect score: 2260

Sequence: 1 NDVARGIVKADVAQNNGLY.....EVQAVNVPSGQRFSLAIVH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/ECTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------------|
| 1 | 2247 | 99.4 | 433 | 4 | US-09-104-623A-4 |
| 2 | 2247 | 99.4 | 433 | 4 | US-09-019-532-4 |
| 3 | 2247 | 99.4 | 433 | 4 | US-09-338-746-4 |
| 4 | 2247 | 99.4 | 433 | 2 | US-08-873-473-43 |
| 5 | 2033 | 90.0 | 641 | 2 | US-08-873-473-42 |
| 6 | 1991.5 | 88.1 | 639 | 4 | US-09-509-814A-4 |
| 7 | 1988.5 | 88.0 | 640 | 4 | US-09-509-814A-8 |
| 8 | 1987.5 | 87.9 | 640 | 4 | US-09-509-814A-6 |
| 9 | 1941.5 | 85.9 | 639 | 4 | US-09-509-814A-1 |
| 10 | 1941.5 | 85.9 | 640 | 4 | US-09-509-814A-2 |
| 11 | 1509 | 66.8 | 345 | 4 | US-09-512-251A-10 |
| 12 | 1509 | 66.8 | 345 | 4 | US-09-515-150A-10 |
| 13 | 1509 | 66.8 | 345 | 4 | US-09-196-281-13 |
| 14 | 438 | 19.4 | 859 | 3 | US-08-894-818B-1 |
| 15 | 438 | 19.4 | 859 | 4 | US-09-445-472-12 |
| 16 | 416.5 | 18.4 | 412 | 4 | US-09-445-472-1 |
| 17 | 416.5 | 18.4 | 412 | 3 | US-08-894-818B-3 |
| 18 | 416.5 | 18.4 | 522 | 4 | US-09-445-472-4 |
| 19 | 416.5 | 18.4 | 654 | 3 | US-08-894-818B-35 |
| 20 | 416.5 | 18.4 | 654 | 4 | US-09-445-472-16 |
| 21 | 395.5 | 17.5 | 659 | 3 | US-08-894-818B-5 |
| 22 | 337 | 14.9 | 520 | 3 | US-09-000-016-7 |
| 23 | 337 | 14.9 | 520 | 4 | US-09-514-340-7 |
| 24 | 337 | 14.9 | 734 | 3 | US-09-000-016-4 |
| 25 | 337 | 14.9 | 734 | 4 | US-09-514-340-4 |
| 26 | 337 | 14.9 | 823 | 3 | US-09-000-016-2 |
| 27 | 337 | 14.9 | 823 | 4 | US-09-514-340-2 |

| | | | | | | |
|----|-------|------|------|---|-------------------|--------------------|
| 28 | 313 | 13.8 | 903 | 1 | US-08-750-532-1 | Sequence 1, Appli |
| 29 | 313 | 13.8 | 1398 | 1 | US-08-750-532-9 | Sequence 9, Appli |
| 30 | 313 | 13.8 | 1398 | 3 | US-08-894-818B-8 | Sequence 8, Appli |
| 31 | 313 | 13.8 | 1398 | 4 | US-09-445-472-6 | Sequence 6, Appli |
| 32 | 280.5 | 12.4 | 237 | 1 | US-08-750-532-18 | Sequence 18, Appli |
| 33 | 273 | 12.1 | 275 | 2 | US-08-750-406A-1 | Sequence 1, Appli |
| 34 | 273 | 12.1 | 275 | 3 | US-09-327-118-1 | Sequence 1, Appli |
| 35 | 270 | 11.9 | 269 | 1 | US-08-431-387-5 | Sequence 5, Appli |
| 36 | 270 | 11.9 | 269 | 1 | US-08-431-387-6 | Sequence 6, Appli |
| 37 | 270 | 11.9 | 269 | 1 | US-08-322-677A-10 | Sequence 10, Appli |
| 38 | 270 | 11.9 | 269 | 1 | US-08-322-676-10 | Sequence 10, Appli |
| 39 | 270 | 11.9 | 269 | 2 | US-08-140-083A-10 | Sequence 10, Appli |
| 40 | 270 | 11.9 | 269 | 3 | US-08-898-218-10 | Sequence 10, Appli |
| 41 | 270 | 11.9 | 269 | 3 | US-08-848-793-10 | Sequence 5, Appli |
| 42 | 270 | 11.9 | 269 | 3 | US-09-255-502-5 | Sequence 5, Appli |
| 43 | 270 | 11.9 | 269 | 3 | US-09-024-532-3 | Sequence 3, Appli |
| 44 | 270 | 11.9 | 269 | 3 | US-08-269-050-4 | Sequence 4, Appli |
| 45 | 270 | 11.9 | 269 | 3 | US-08-090-207-1 | Sequence 1, Appli |

ALIGNMENTS

RESULT 1
US-09-104-623A-4
; Sequence 4, Application US/09104623A
; Patent No. 6303752
; GENERAL INFORMATION:
; APPLICANT: Olsen, Arne Agerlin
; APPLICANT: Fatum, Tine Muxoll
; APPLICANT: Roggen, Heinz-Josef
; APPLICANT: Roggen, Erwin Ludo
; TITLE OF INVENTION: A Modified Polypeptide
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6303752o No. 6303752disk of No. 6303752th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,623A
; FILING DATE: 25-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 5256.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; STRAIN: Bacillus sp. Y
US-09-104-623A-4

Query Match 99.4%; Score 2247; DB 4; Length 433;

Best Local Similarity 99.3%; Pred. No. 4.4e-169;

Matches 430; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGLYGQQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRIN 60


```
Db 1 NDVARGIVKADVAQNNGYGLGQQLVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
Qy 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIIMDSGGGLGGLPSNLTLFSQAWNA 120
Db 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIIMDSGGGLGGLPSNLTLFSQAWNA 120
Qy 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Db 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Qy 181 VGATENYRPSFGSIADPNHIAQFSRSGATRDGRIPKDVTAAGTFTILSARSSSLAPDSSFW 240
Db 181 VGATENYRPSFGSIADPNHIAQFSRSGATRDGRIPKDVTAAGTFTILSARSSSLAPDSSFW 240
Qy 241 ANYNSKYAYMGTSMTPTIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
Db 241 ANYNSKYAYMGTSMTPTIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
Qy 301 NGDQGWGRVTLNKSUNVAVNEATATLQOKATYSFOAQAGKPLKISLWVTDAPGSTTAS 360
Db 301 NGDQGWGRVTLNKSUNVAVNEATATLQOKATYSFOAQAGKPLKISLWVTDAPGSTTAS 360
Qy 361 YTLVNDLDLVTAPNGQKYGNDFSYPYDNNWDGRNNVENFINAPQSGTYIIIEVQAYNV 420
Db 361 YTLVNDLDLVTAPNGQKYGNDFSYPYDNNWDGRNNVENFINAPQSGTYIIIEVQAYNV 420
Qy 421 PSGPQRFSLAIVH 433
Db 421 PSGPQRFSLAIVH 433
```

RESULT 2

US-09-019-532-4

; Sequence 4, Application US/09019532B

; Patent No. 6416756

; GENERAL INFORMATION:

; APPLICANT: Olsen, Arne Agerlin

; APPLICANT: Olsen, Arnette

; TITLE OF INVENTION: A Modified Enzyme for Skin Care

; FILE REFERENCE: 4922,204-US

; CURRENT APPLICATION NUMBER: US/09/019,532B

; EARLIER FILING DATE: 1998-02-05

; EARLIER APPLICATION NUMBER: 0038/97

; EARLIER FILING DATE: 1997-01-10

; EARLIER APPLICATION NUMBER: 0754/97

; EARLIER FILING DATE: 1997-06-25

; EARLIER APPLICATION NUMBER: 60/051,381

; EARLIER FILING DATE: 1997-07-07

; EARLIER APPLICATION NUMBER: PCT/DK98/000015

; EARLIER FILING DATE: 1998-01-12

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 4

; LENGTH: 433

; TYPE: PRT

; ORGANISM: Bacillus sp.

US-09-019-532-4

Query Match

Best Local Similarity 99.4%; Score 2247; DB 4; Length 433;

Matches 430; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 NDVARGIVKADVAQNNGYGLGQQLVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
Db 1 NDVARGIVKADVAQNNGYGLGQQLVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
Qy 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIIMDSGGGLGGLPSNLTLFSQAWNA 120
Db 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIIMDSGGGLGGLPSNLTLFSQAWNA 120
Qy 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
```

Db 301 SGDQGWGRVTLKSLNVAAYNEATATGQKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
Qy 361 YTLVNDLDLVTAPNGQKYGNDPSYPYDNNWNGRNNVNFVNAPOSGTYIIIEVQAYNV 420
Db 361 YTLVNDLDLVTAPNGQKYGNDPSYPYDNNWNGRNNVNFVNAPOSGTYIIIEVQAYNV 420
Qy 421 PSGQRFSLAIVH 433
Db 421 PSGQRFSLAIVH 433

RESULT 4
US-08-873-479-43
; Sequence 43, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Lynne, Christianson
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58917010 No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: US/08/873,479
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-873-479-43

Query Match 99.4%; Score 2247; DB 2; Length 635;
Best Local Similarity 99.3%; Pred. No. 7,5e-169;
Matches 430; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NDVARGIVKADVQNNYGLYGQGVAVADTGLDTRNDSSNHEAFRGKITIYALGRTN 60
Db 203 NDVARGIVKADVQNNYGLYGQGVAVADTGLDTRNDSSNHEAFRGKITIYALGRTN 262

Qy 61 NASDPNGHGTTHVAGSVLGNALNKGAPOANLVFQSIMDSSGGLGSLPSNLTLFSQAWNA 120
Db 263 NASDPNGHGTTHVAGSVLGNALNKGAPOANLVFQSIMDSSGGLGSLPSNLTLFSQAWNA 322

Qy 121 GARIHTNSWGAFTNGAYTANSRQVDEYVRNNDMTVLFPAAGNEGNSGTISAPGTAKNAIT 180
Db 323 GARIHTNSWGAFTNGAYTANSRQVDEYVRNNDMTVLFPAAGNEGNSGTISAPGTAKNAIT 382

Qy 181 VCATENYRPSFGSIADNPNHIAQFSSRGATRDGRIPKPDVTAPGTFFLSARSLAPDSFW 240
Db 383 VCATENYRPSFGSIADNPNHIAQFSSRGATRDGRIPKPDVTAPGTFFLSARSLAPDSFW 442

Qy 241 ANYNSKIAYMGTSMTATPIVAGNVAQLREHFIKRGITPKPSLIKAALIAGATDVGLGYP 300
Db 443 ANYNSKIAYMGTSMTATPIVAGNVAQLREHFIKRGITPKPSLIKAALIAGATDVGLGYP 502
Qy 301 NDDQGWGRVTLKSLNVAAYNEATATGQKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
Db 503 SGDQGWGRVTLKSLNVAAYNEATATGQKATYSFOAQAGKPLKISLVWTDAPGSTTAS 562
Qy 361 YTLVNDLDLVTAPNGQKYGNDPSYPYDNNWNGRNNVNFVNAPOSGTYIIIEVQAYNV 420
Db 563 YTLVNDLDLVTAPNGQKYGNDPSYPYDNNWNGRNNVNFVNAPOSGTYIIIEVQAYNV 622

Qy 421 PSGQRFSLAIVH 433
Db 623 PSGQRFSLAIVH 635

RESULT 5
US-08-873-479-42
; Sequence 42, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Lynne, Christianson
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; TITLE OF INVENTION: Having Protease Activity
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58917010 No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: US/08/873,479
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-873-479-42

Query Match 90.0%; Score 2033; DB 2; Length 641;
Best Local Similarity 88.7%; Pred. No. 5,5e-152;
Matches 384; Conservative 25; Mismatches 24; Indels 0; Gaps 0;

Qy 1 NDVARGIVKADVQNNYGLYGQGVAVADTGLDTRNDSSNHEAFRGKITIYALGRTN 60
Db 209 NDVARGIVKADVQNNYGLYGQGVAVADTGLDTRNDSSNHEAFRGKITIYALGRTN 268

Qy 61 NASDPNGHGTTHVAGSVLGNALNKGAPOANLVFQSIMDSSGGLGSLPSNLTLFSQAWNA 120
Db 269 NASDPNGHGTTHVAGSVLGNALNKGAPOANLVFQSIMDSSGGLGSLPSNLTLFSQAWNA 328

QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTGISAPGTAKNAIT 180
 Db 329 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTGISAPGTAKNAIT 388
 QY 181 VGATENLRPSFGSIADNPNHIAQFSSRGATDGRKPDVTPAGTFFILSARSLAPDSF 240
 Db 389 VGATENLRPSFGSIADNPNHIAQFSSRGATDGRKPDVTPAGTFFILSARSLAPDSF 448
 QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRGTIPKPSLIKAALIAGATDVGLGY 300
 Db 449 ANHDSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRGTIPKPSLIKAALIAGATDVGLGY 508
 QY 301 NGDQGWGRVTLNKSUNVAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTA 360
 Db 509 NGDQGWGRVTLNKSUNVAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTA 568
 QY 361 YTLVNDLDLIVITAPNGQKTVGNDFSPYDNNWGDGRNNVFNINAPQSGTYIIIEVQAYN 420
 Db 569 LTLVNDLDLIVITAPNGQKTVGNDFSPYDNNWGDGRNNVFNINAPQSGTYIIIEVQAYN 628
 QY 421 PSQPORFSLAIVH 433
 Db 629 PVPQPNFSLAIVH 641

RESULT 6

US-09-509-814A-4
 ; Sequence 4, Application US/09509814A
 ; Patent No. 6376227

; GENERAL INFORMATION:
 ; APPLICANT: TAKAIWA, MIKIO

; APPLICANT: OKUDA, MITSUYOSHI

; APPLICANT: SAEKI, KATSUHISA

; APPLICANT: KUBOTA, HIROMI

; APPLICANT: HITOMI, JUN

; APPLICANT: KAGEYAMA, YASUSHI

; APPLICANT: SHIKATA, SHITSUM

; APPLICANT: NOMURA, MASAFUMI

; TITLE OF INVENTION: ALKALINE PROTEASE

; FILE REFERENCE: 0327-0832-OPCT

; CURRENT APPLICATION NUMBER: US/09/509,814A

; CURRENT FILING DATE: 2000-04-06

; PRIOR APPLICATION NUMBER: PCT/JP98/04528

; PRIOR FILING DATE: 1998-10-07

; PRIOR APPLICATION NUMBER: JP 9-274570

; PRIOR FILING DATE: 1997-06-08

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 4

; LENGTH: 639

; TYPE: PRT

; ORGANISM: Bacillus sp.

US-09-509-814A-4

Query Match 88.1%; Score 1991.5; DB 4; Length 639;

Best Local Similarity 87.8%; Pred. No. 1e-148;

Matches 381; Conservative 27; Mismatches 25; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNGYLGQGVAVADTGLDTRNDSSMHEAPRGKITAYALGRTN 60
 Db 206 NDVARGIVKADVAQNNGYLGQGVAVADTGLDTRNDSSMHEAPRGKITAYALGRTN 265
 QY 61 NASDPNGHGHVAGSVLGN-ALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWN 119
 Db 266 NANDTNGHGHVAGSVLGNATNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAFS 325
 QY 120 AGARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTGISAPGTAKNAI 179
 Db 326 AGARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTGISAPGTAKNAI 385
 QY 180 TVGATENLRPSFGSIADNPNHIAQFSSRGATDGRKPDVTPAGTFFILSARSLAPDSF 239

Db 386 TVGATENLRPSFGSIADNPNHIAQFSSRGATDGRKPDVTPAGTFFILSARSLAPDSF 445
 QY 240 WANYNSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRGTIPKPSLIKAALIAGATDVGLGY 299
 Db 446 WANHDSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRGTIPKPSLIKAALIAGATDVGLGY 505
 QY 300 PNGDQGWGRVTLNKSUNVAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTA 359
 Db 506 PNGDQGWGRVTLNKSUNVAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTA 565
 QY 360 SYTLVNDLDLIVITAPNGQKTVGNDFSPYDNNWGDGRNNVFNINAPQSGTYIIIEVQAYN 419
 Db 566 SYTLVNDLDLIVITAPNGQKTVGNDFSPYDNNWGDGRNNVFNINAPQSGTYIIIEVQAYN 625
 QY 420 VPSQPORFSLAIVH 433
 Db 626 PVPQPNFSLAIVH 639

RESULT 7

US-09-509-814A-8

; Sequence 8, Application US/09509814A

; Patent No. 6376227

; GENERAL INFORMATION:

; APPLICANT: TAKAIWA, MIKIO

; APPLICANT: OKUDA, MITSUYOSHI

; APPLICANT: SAEKI, KATSUHISA

; APPLICANT: KUBOTA, HIROMI

; APPLICANT: HITOMI, JUN

; APPLICANT: KAGEYAMA, YASUSHI

; APPLICANT: SHIKATA, SHITSUM

; APPLICANT: NOMURA, MASAFUMI

; TITLE OF INVENTION: ALKALINE PROTEASE

; FILE REFERENCE: 0327-0832-OPCT

; CURRENT APPLICATION NUMBER: US/09/509,814A

; CURRENT FILING DATE: 2000-04-06

; PRIOR APPLICATION NUMBER: PCT/JP98/04528

; PRIOR FILING DATE: 1998-10-07

; PRIOR APPLICATION NUMBER: JP 9-274570

; PRIOR FILING DATE: 1997-06-08

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 8

; LENGTH: 640

; TYPE: PRT

; ORGANISM: Bacillus sp.

US-09-509-814A-8

Query Match 88.0%; Score 1988.5; DB 4; Length 640;

Best Local Similarity 87.3%; Pred. No. 1.7e-148;

Matches 379; Conservative 29; Mismatches 25; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNGYLGQGVAVADTGLDTRNDSSMHEAPRGKITAYALGRTN 60
 Db 207 NDVARGIVKADVAQNNGYLGQGVAVADTGLDTRNDSSMHEAPRGKITAYALGRTN 266
 QY 61 NASDPNGHGHVAGSVLGN-ALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWN 119
 Db 267 NANDTNGHGHVAGSVLGNSTNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAFS 326
 QY 120 AGARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTGISAPGTAKNAI 179
 Db 327 AGARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTGISAPGTAKNAI 386
 QY 180 TVGATENLRPSFGSIADNPNHIAQFSSRGATDGRKPDVTPAGTFFILSARSLAPDSF 239
 Db 387 TVGATENLRPSFGSIADNPNHIAQFSSRGATDGRKPDVTPAGTFFILSARSLAPDSF 446
 QY 240 WANYNSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRGTIPKPSLIKAALIAGATDVGLGY 299
 Db 447 WANHDSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRGTIPKPSLIKAALIAGATDVGLGY 506
 QY 300 PNGDQGWGRVTLNKSUNVAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTA 359

```
Db 507 PNGNQGRVTLDKSLNVAAYNESSLSSTOKATYSFTATAGKPLKISLVMSDAPASTTA 566
Qy 360 SYTLVNDLDLVTAPNGQKVGNDPSYDNNWGRNNVFNAPQSGTYIIIEVQAYN 419
Db 567 SVTLVNDLDLVTAPNGTYVGNDFTSYDNNWGRNNVFNAPQSGTYIIIEVQAYN 626
Qy 420 VPSGPQRFSLAIVH 433
Db 627 VEVGPQNFSLAIVN 640

RESULT 8
US-09-509-814A-6
; Sequence 6, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUW
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-OPCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
; NAME/KEY: misc feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (46)..(46)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (70)..(70)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (74)..(74)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (89)..(89)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (102)..(102)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (105)..(105)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (128)..(128)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (130)..(130)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature

Query Match 87.9%; Score 1987.5; DB 4; Length 640;
Best Local Similarity 87.3%; Pred. No. 2.1e-148;
Matches 379; Conservative 29; Mismatches 25; Indels 1; Gaps 1;
Qy 1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALVALGRTN 60
Db 207 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALVALGRTN 266
Qy 61 NASDPNGHGHVAGSVLGN-ALNKGMAPOANLVFQSIMDSGGLGGLPSNLNTLFSQAWN 119
Db 267 NANTNGHGHVAGSVLGNSTKNGMAPOANLVFQSIMDSGGLGGLPSNLNTLFSQAYS 326
Qy 120 AGARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLPAAAGNEGPNSTGISAPGTAKNAI 179
Db 327 AGARIHTNSGAPVNGAYTDSRNVDYRKDMTILPAAAGNEGPNSTGISAPGTAKNAI 386
Qy 180 TVGATENYRPSFGSIADNPNHIAQSSRGATDRGKIPDVTAPGTFTLSARSSLAPDSFP 239
Db 387 TVGATENLRFSGYADNHNHVAQSSRGFTDKGRIPKPDVMAPTFTLSARSSLAPDSFP 446
Qy 240 WANNYSKYAMGTSMTATPVAGNVQALREHPKRGITPKPSLTKAALIAGATDVGLGY 299
Db 447 WANHSKYAMGTSMTATPVAGNVQALREHFVKGRTIPKPSLTKAALIAGADIGLGY 506
Qy 300 PNGDQGRVTLNKLNLVAAYNEATALATGQKATYSFOAQKPLKISLVWTDAPGSTTA 359
Db 507 PNGNQGRVTLDKSLNVAAYNESSLSSTOKATYSFTATAGKPLKISLVMSDAPASTTA 566
Qy 360 SYTLVNDLDLVTAPNGQKVGNDPSYDNNWGRNNVFNAPQSGTYIIIEVQAYN 419
Db 567 SVTLVNDLDLVTAPNGTYVGNDFTSYDNNWGRNNVFNAPQSGTYIIIEVQAYN 626
```

```
Qy 420 VPSGPQRFSLAIVH 433
Db 627 VEVGPQNFSLAIVN 640

RESULT 9
US-09-509-814A-1
; Sequence 1, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUW
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-OPCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.
; NAME/KEY: misc feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (46)..(46)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (70)..(70)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (74)..(74)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (89)..(89)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (102)..(102)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (105)..(105)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (128)..(128)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
; LOCATION: (130)..(130)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc feature
```

LOCATION: (131)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (594)..(594)

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1

Query Match 85.9%; Score 1941.5; DB 4; Length 639;

Best Local Similarity 86.2%; Pred. No. 8.8e-145;

Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALYALGRIN 60
DB 206 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALYALGRIN 265
QY 61 NASDPNGHGHVAGSVLGN-ALNKGVAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 119
DB 266 NANTNGHGHVAGSVLGNKGTNKGVAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAXS 325
QY 120 AGARHTNSWGA PVNGAYTANSRQYDEVYRNDMTVLPAAGNEGPNSTGISAPGTAKNAI 179
DB 326 AGARHTNSWGA PVNGAYTANSRQYDEVYRNDMTVLPAAGNEGPNSTGISAPGTAKNAI 385
QY 180 TVGATENYRPSFGSIADPNPHIAQFSSRGATRDGRIKPDVTAPGTFFILSARSLAPDSGF 239
DB 386 TVGATENLRPSFGSYADNINHVAQFSSRGATKDGRIKPDVMAPTGTXILSARSLAPDSF 445
QY 240 WANYSKYAYMGTSMATPIVAGNVAQLREHFKRGITPKSLIKAALIAGATDVGLGY 299
DB 446 WANHDSKYAYMGTSMATPIVAGNVAQLREHFKRGITPKSLIKAALIAGAADXGLGY 505
QY 300 PNGDQGWGRVTLNKS LNVA YNEATALATGQKATYSFOAQAGPKLKI SLVWTDAPGSTTA 359
DB 506 PNGNQGWRVTLNKS LNVA YNEATALATGQKATYSFOAQAGPKLKI SLVWTDAPGSTTA 565
QY 360 SYTLVNDLVLITAPNGQKQYGVNDPSYPYDNNWDCGRNNVNFINAPQSGTVIIISVQAVN 419
DB 566 SVTLVNDLVLITAPNGTYGVNDPFXKXXNWDGRNNVNFINAPQSGTVIIISVQAVN 625
QY 420 VPSGPQRFSLAIHV 433
DB 626 VPVGPQXPSLAIVN 639

RESULT 10

US-09-509-814A-2

Sequence 2, Application US/09509814A

Patent No. 6376227

GENERAL INFORMATION:

APPLICANT: TAKAIWA, MIKIO

APPLICANT: OKUDA, MITSUYOSHI

APPLICANT: SAEKI, KATSUHIISA

APPLICANT: KUBOTA, HIROMI

APPLICANT: HITOMI, JUN

APPLICANT: KAGEYAMA, YASUSHI

APPLICANT: SHIKATA, SHITSUM

APPLICANT: NOMURA, MASAFUMI

TITLE OF INVENTION: ALKALINE PROTEASE

FILE REFERENCE: 0327-0832-0PCT

CURRENT APPLICATION NUMBER: US/09/509,814A

CURRENT FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: PCT/JP98/04528

PRIOR FILING DATE: 1998-10-07

PRIOR APPLICATION NUMBER: JP 9-274570

PRIOR FILING DATE: 1997-06-08